

МАШИННОЕ ОБУЧЕНИЕ В СФЕРЕ ПЛАТФОРМИЗАЦИИ БИЗНЕС-МОДЕЛЕЙ

И. А. КАРАЧУН¹⁾

¹⁾ Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь

Устойчивая тенденция к платформизации многих сфер экономики и общественной жизни позволяет считать ее ключевой концепцией, которую следует учитывать, когда речь идет о промышленных или, скорее, цифровых революциях. Платформенные экосистемы замещают традиционные модели сначала в ИТ-сфере, а потом и за ее пределами во многом благодаря оцифровке продуктов, услуг и бизнес-процессов. В свою очередь, это требует фундаментального изменения стратегического мышления. Владелец платформы должен сознательно управлять ее эволюцией в рамках определенной модели, позволяющей достичь максимального удобства, скорости операций и разнообразия контрагентов. Рассмотрены процессы внедрения машинного обучения в различные системы финансового моделирования и принятия решений. В работе представлен отраслевой контекст машинного обучения в сфере финансов, обсуждаются важные события, сформировавшие потребность финансовой отрасли в машинном обучении, и уникальные барьеры для внедрения машинного обучения в сферу финансов. Финансовая индустрия по-разному приняла машинное обучение. Некоторые ключевые примеры демонстрируют природу машинного обучения и специфику его практического применения.

Ключевые слова: машинное обучение; краудтехнологии; платформа; бизнес-модель; финансовые технологии.

MACHINE LEARNING IN THE FIELD OF BUSINESS MODEL PLATFORMISATION

I. A. KARACHUN^a

^a Belarusian State University, 4 Niezaliežnasci Avenue, Minsk 220030, Belarus

The strong trend towards platformisation in many areas of the economy and public life allows us to consider it a fundamental concept known and understood for industrial or somewhat digital revolutions. Platform ecosystems are replacing traditional models, first in the IT field and then beyond, mainly due to the digitisation of products, services and business processes. In turn, this requires a fundamental change in strategic thinking. The platform owner must consciously manage its evolution within a specific model, allowing maximum convenience, speed of operations, and diversity of counterparties. We examined the processes of implementing machine learning in various financial modelling and decision-making systems. This paper introduces the industry context for machine learning in finance, discusses essential developments that have shaped the need for machine learning in the finance industry, and discusses unique barriers to adoption. The financial industry has embraced machine learning to varying degrees of sophistication. Some key examples demonstrate the nature of machine learning and how it is used in practice.

Keywords: machine learning; crowd technologies; platform; business model; financial technology.

Образец цитирования:

Карачун ИА. Машинное обучение в сфере платформизации бизнес-моделей. *Журнал Белорусского государственного университета. Экономика*. 2022;1:79–88.

For citation:

Karachun IA. Machine learning in the field of business model platformisation. *Journal of the Belarusian State University. Economics*. 2022;1:79–88. Russian.

Автор:

Ирина Андреевна Карачун – кандидат экономических наук, доцент; заведующий кафедрой цифровой экономики экономического факультета.

Author:

Irina A. Karachun, PhD (economics), docent; head of the department of digital economy, faculty of economics.
karachun@bsu.by
<https://orcid.org/0000-0002-0132-5064>

В последнее десятилетие была наглядно продемонстрирована способность цифровых технологий изменять методы работы организаций. Сейчас цифровизация вышла за рамки улучшения бизнес-процессов, она ставит под сомнение причины существования компаний и их фундаментальную ценность. Явление цифрового разрушения, или сбоя (*digital disruption*) [1], развивается и становится реальной угрозой.

Концепция цифрового разрушения понимается как совокупность компонентов, которые образуют серию явлений, постепенно охватывающих все отрасли промышленности и экономическую жизнь [2]. Точный характер и нюансы применения технологической инновации проявляются в культурном и историческом контексте, уникальном для каждой отрасли, поэтому невозможно определить единую траекторию разрушения.

Еще со времен Р. Коуза (1937) [3] фирма определяется как связующее звено сделок для внутреннего управления операциями, сокращения времени и затрат. В цифровую эпоху стоимость внутренней организации операций в фирме часто превышает стоимость рыночных услуг, вследствие чего выгоды от организации меняются на противоположные, а структура организации операций становится неоптимальной и неустойчивой. Например, доступ к облачным вычислительным службам сводит на нет необходимость создания собственных вычислительных сред, поскольку рыночные операции дешевле внутреннего предоставления услуг – создания собственного ИТ-отдела. Под действием законов Мура [4] и Меткалфа [5] стоимость вычислений с доступом к интернету снижается. Делать прогнозы относительно оптимального баланса собственного производства и аутсорсинга невозможно, так как окружающая среда постоянно меняется. Очень сложно предсказать, какой степени дезагрегации подвергнутся фирмы и в каком объеме останутся существовать организации. Методом проб и ошибок предприниматели пытаются преодолеть неэффективность масштаба за счет экономии на операциях и снижении затрат, определяя оптимальную структуру [6]. Каждая отрасль имеет собственный поток создания ценности, поэтому не все отрасли или компании будут затронуты одинаково и кто-то в той или иной степени выиграет от снижения операционных издержек.

Исследование базируется на работах по двусторонним рынкам начиная с публикаций Ж.-Ш. Роше, Ж. Тироль [7] и М. Армстронга [8]. В рамках довольно обширного перечня литературы по платформенной экономике проанализированы исследования связи между платформенными структурами, поставщиками контента и комплементорами. Например, Д. Йофи и М. Квак [9] фокусируются на конфликтах между платформами и комплементорами, А. Хаджиу [10] подчеркивает эндогенный характер сетевых эффектов внутри платформ, возникающих в результате взаимодействия поставщиков контента и потребителей. А. Галеотти и Ж. Морага-Гонзалес [11], изучавшие двустороннюю платформу, которая привлекает поставщиков контента и потребителей, отмечали, что увеличение ассортимента повышает ценность платформы для потребителей, но ослабляет конкуренцию между поставщиками, тем самым создавая стимулы для платформы повышать плату для обеих сторон. На основании анализа данных, собранных в США, С. Сенамо и Ж. Сантало [12] выявили, что, помимо использования преимуществ положительных сетевых эффектов, связанных с доминированием на рынке, платформам необходимо успешно управлять стимулами для поставщиков контента и активировать рост экосистемы. В свою очередь, Дж. Паркер и М. ван Алстин [13] сосредоточились на оптимальном уровне открытости платформы, а также на длительности предоставления прав интеллектуальной собственности поставщиков экосистеме платформы. Автор настоящей статьи обобщает эти исследования, распространяет их результаты на финансовые технологии, а также сосредоточивает внимание на разрушительных аспектах цифровизации, трансформации производственно-технологических моделей и бизнес-моделей, изменениях потребительских ожиданий и поведения.

В настоящее время почти все отрасли экономики претерпевают фундаментальные преобразования, основанные на принципе, согласно которому все, что может быть оцифровано, будет оцифровано. Это, в свою очередь, ведет к цифровизации всего мира, так как программное обеспечение становится ключевым фактором успеха во многих сферах бизнеса, а изменения кажутся скорее разрушением старого порядка, чем эволюцией. В качестве основной концепции и бизнес-модели новой вехи экономического развития выступает платформизация, практическими примерами которой являются такие компании, как *Airbnb*, *Facebook*, *LinkedIn*, *Uber*, *Яндекс* и *Alibaba Group*, при этом в теории управления модель платформы все еще находится на ранней стадии обсуждения и систематизации.

Основываясь на анализе предшествующих исследований, можно отметить, что при рассмотрении цифровых изменений базовой является концепция экосистемы. Этот заимствованный из биологии термин обычно относят к группе агентов, зависящих от деятельности друг друга [14]. Ученые подчеркивают различные аспекты этой структуры, которые условно можно разделить на несколько направлений. В центре бизнес-экосистемы находятся предприятие и его окружающая среда. Инновационная экосистема сосредоточена вокруг некоей инновации и совокупности субъектов, которые ее

поддерживают. Экосистема платформы рассматривает, как агенты организуются вокруг платформы. Третье направление исследований посвящено определенному классу технологических решений (платформам) и взаимозависимостям между спонсорами, клиентами и комплементорами платформ. С бытующей точки зрения экосистема включает спонсора платформы и всех поставщиков комплементарных товаров и услуг, которые делают платформу более ценной для потребителей [15; 16]. Экосистема платформы принимает форму колеса (*hub and spoke*) – множества периферийных фирм, подключенных к центральной платформе с помощью общих или открытых технологий и (или) технических стандартов (которые могут быть интерфейсами программирования или комплектами для разработки программного обеспечения). Подключаясь к платформе, комплементоры могут не только генерировать дополнительные инновации, но и получать прямой или косвенный доступ к клиентам платформы (как, например, независимые поставщики программного обеспечения, аффилированные с SAP S/4HANA [15], или разработчики видеоигр для различных консолей [12]). Соответственно, экосистемы платформы рассматриваются как полурегулируемые рынки, способствующие предпринимательской деятельности и координируемые спонсором платформы [17], или как многосторонние рынки, позволяющие осуществлять транзакции между различными группами пользователей [12]. В свою очередь, автор настоящей статьи полагает, что разделить клиентов платформы и комплементоров практически невозможно, поскольку они, как правило, являются частями одной структуры, объединенными как технологией, так и бизнес-моделью. Следовательно, платформа тождественна экосистеме, и дальнейшие разработки целесообразно проводить в направлении идентификации цифровой платформенной бизнес-модели и образующих ее акторов.

Следует отметить, что благодаря платформизации появилась возможность заменять услуги или продукты на более качественные без повышения цены или дополнительных затрат на переключение. Для традиционных фирм угроза исходит от краудсорсинга и способности реагировать на предполагаемый неудовлетворенный спрос без фиксированных затрат. Это позволяет экономить и эффективно обслуживать сегмент рынка, а также полагаться на поставщиков услуг для удовлетворения потребностей остальной части рынка. Так, например, в случае экспертных посреднических услуг, предоставляемых туристическими агентствами, это оказалось недостаточным, поскольку благодаря интернету потребители имели доступ к информации, ранее доступной только через специальное расписание рейсов авиакомпаний. Стремление авиакомпаний близко узнать своего клиента для эффективного самообслуживания в совокупности привело к удалению туристических агентов из звеньев цепочки создания стоимости.

Склонность к разрушению отрасли во многом зависит от ее структуры и производительности, степени сложности цифровых технологий и доступности информации для агрегации. Открытость информации, которой можно обмениваться в цифровом виде, в сочетании со сжатием цепочки создания стоимости значительно увеличивает степень потенциального нарушения, при этом существуют барьеры, защищающие и отрасль, и сотрудников, – препятствия, с которыми сталкиваются фирмы, когда хотят войти в отрасль, расширить или покинуть ее. Считается, что они снижают конкурентоспособность промышленного сектора, поскольку препятствуют свободному входу в него и выходу из него, обеспечивая фирмы защитным барьером от разрушительных инноваций.

Структурные барьеры состоят из целого ряда препятствий, которые включают регулирование, доступ к капиталу (финансовому или интеллектуальному), масштабы деятельности или возможность доступа к цепочкам поставок, хотя цифровизация разрушает многие из них.

Для любой фирмы, желающей превратиться в крупного поставщика, важен доступ к финансовому капиталу. Эффективность рынков капитала обеспечивает стартапам возможность получать средства, необходимые для роста и достижения критического операционного порога, чтобы конкурировать с существующими поставщиками. В свою очередь, стоимость венчурных фондов, как правило, относительно высока, что является сдерживающим фактором для собственников и команд небольших, но динамичных стартапов. В результате доля на рынке многих потенциально разрушающих отрасль компаний остается меньшей, чем могла бы быть. Такие финансовые инновации, как краудсорсинг, сочетаются с низкими стартовыми затратами благодаря возможностям совместного использования, например, облачных вычислений (*software as a service* – программное обеспечение как услуга), что позволяет цифровым стартапам вырасти до конкурентных масштабов в значительно меньший срок.

Возможность достижения конкурентоспособного операционного масштаба тесно связана с доступом к капиталу. Необходимость амортизации затрат при достаточно большом объеме продаж может иметь решающее значение, когда требуются большие инвестиции в производство, персонал или развитие рынка. В цифровом ландшафте традиционная экономика за счет масштабов производства быстро становится неэффективной, поскольку с учетом инвестиций в инфраструктуру цепочки поставок или с участием позиций продукта сопутствующее отсутствие гибкости создает барьер для выхода из бизнес-модели, линейки продуктов или сферы предоставления услуг.

Доступ к контрактным услугам, экспертам по предмету и профессионалам стал практически бесшовным благодаря возможностям поиска в интернете и совместной работы, поэтому доступ к специализированным навыкам больше не является таким серьезным барьером, как раньше. Это позволяет небольшим компаниям сосредоточиться только на своих ключевых компетенциях и заключать контракты на целые функциональные области, которые могут включать бухгалтерский учет, маркетинг, производство, управление персоналом, разработку программного обеспечения и управление проектами.

Государственное регулирование и законодательство могут создавать серьезные препятствия для проникновения цифровых компаний в отрасль, поскольку существующие группы лоббируют вмешательство правительства для сдерживания изменений рынка. Наиболее серьезным препятствием является патентное право, которое крупные фирмы используют не только для законной защиты интеллектуальной собственности, но и для того, чтобы помешать конкурентам и стартапам оспаривать их потоки доходов. Использование судебных разбирательств также можно рассматривать как самостоятельное препятствие для входа компании в отрасль.

Возможно, самым значительным барьером для входа в цифровую эпоху станет лояльность клиентов или, скорее, ее отсутствие. Способность защищать интересы клиентов будет самым сильным фактором, определяющим успех любого продукта или услуги. Задержка передачи знаний уже практически отсутствует, новые услуги, продукты, цены, дизайн, функции и мнения доступны в течение нескольких минут после изменения рынка. Цифровые технологии идеально подходят для вырабатывания лояльности благодаря способности прямо и косвенно собирать информацию о клиентах, что требует более широкого использования аналитики и больших данных. Построение алгоритмических отношений станет ключевым компонентом стратегий лояльности клиентов, а сложное и творческое использование данных – важнейшей компетенцией. В настоящее время такие возможности являются прерогативой крупных организаций, с которыми сотрудничают лишь некоторые технологически специализированные фирмы.

Расходы на повышение узнаваемости и ценности бренда – существенный сдерживающий фактор для инновационных компаний. Получить доверие потребителя чрезвычайно дорого и требует постоянных инвестиций. Социальные сети сейчас предоставляют экономичный канал для менее обеспеченных компаний, использующих сарафанное радио и сетевые эффекты для распространения информации о продуктах. Но это может продлиться недолго, поскольку рекламная индустрия стремится занять данную сферу, чтобы защитить свои потоки доходов. Усложнение и персонализация рекламы поставят стартапы в невыгодное для их роста положение и могут увеличить сопротивление инновациям.

К клиентоориентированным барьерам можно отнести и затраты на переключение. Чем удобнее клиенту пользоваться продуктом или услугой, тем большее психологическое усилие ему понадобится, чтобы изменить свои привычки. Простота использования, внедрения и интеграции оказывается решающим фактором для снижения затрат на переключение. Часто первоначальный продукт не удерживается на рынке, хотя и служит для создания нового рыночного пространства и открытия сознания потребителей для новой формы предоставления услуги. В данном случае примером могут служить устройства *BlackBerry*, которые сначала заняли рынок мобильных телефонов для бизнеса, но уступили свои позиции смартфонам *iPhone* компании *Apple*, а затем телефонам с операционной системой *Android*.

Существуют также организационные, личные и управленческие барьеры. Такие мягкие барьеры включают культуру, привычки и эвристику, а также способности манипулировать цифровыми инструментами. Например, способность внедрять технологию связана с ограниченностью интеллектуальных ресурсов: необходимость понимания очень сложных технологий создает барьер для получения знаний, поскольку потенциальным участникам требуется дополнительное обучение. Цифровые технологии широко распространяются благодаря постоянному сокращению затрат на компьютеры и создание сетей, тем не менее препятствием по-прежнему остается неспособность трансформировать бизнес за счет использования расширенных цифровых знаний. На данный момент производительность в основном повышается благодаря автоматизации, что интуитивно понятно и осязаемо. Многие руководители высшего звена не владеют информационными технологиями, что снижает их способность адаптироваться к внедрению инноваций и создает значительный барьер для этого процесса. Неспособность организаций к изменениям представляет собой проблему для их входа на рынок, поскольку преобладающая операционная культура препятствует цифровому переопределению.

Бизнес-модели и структуры чаще всего разрабатывают академические исследователи и консалтинговые компании, поэтому их продукты не всегда отвечают современным вызовам. Быстрое и дестабилизирующее воздействие цифровых технологий требует различных подходов, учитывающих динамический характер изменений. Операционных моделей, нацеленных на постоянное повышение эффективности цепочки создания стоимости, недостаточно, так как цепочка создания стоимости сама скоро станет пользователем, а не производителем. Ценности и принципы влияют на решения клиентов. В открытом

обществе с доступом к информации во всех социальных секторах фирмы должны публично отвечать за свои действия не только перед акционерами, но и перед всеми заинтересованными сторонами. Непорядочное поведение нельзя долго скрывать, а последствия публичного осуждения разрушительны.

Развитие современных информационных технологий влечет за собой небывалый рост объемов вычислительных ресурсов и больших данных. Благодаря этому методы машинного обучения, доступные через наборы инструментов с открытым исходным кодом, приобретают популярность среди аналитиков и разработчиков. Модели глубокого обучения оказались чрезвычайно успешными в широком диапазоне приложений, включающем обработку изображений, геймификацию обучения, нейробиологию, энергосбережение и медицинскую диагностику.

Машинное обучение в сфере финансов объединяет несколько математических дисциплин: финансовую эконометрику, статистические вычисления, вероятностное и динамическое программирование и даже распознавание образов. Основными потребителями этой технологии стали области финансов, ориентированные на алгоритмы, например алгоритмическая торговля. Ключевой проблемой на пути к пониманию машинного обучения является отсутствие устоявшихся теорий и концепций, которые необходимы для анализа финансовых временных рядов. Существуют множество неправильных представлений и ограниченное понимание возможностей этой области, а эффективные методы машинного обучения остаются плохо изученными и часто математически необоснованными.

Увеличение количества машиночитаемых данных о действиях в финансовой системе в сочетании с постоянным ростом вычислительной мощности и емкости хранилищ имеет важные последствия для каждого аспекта финансового моделирования. Из-за финансового кризиса 2008 г. надзорные органы многих стран стали осуществлять регулирование, основываясь на анализе данных, внедрять программы стресс-тестирования [18]. Все более важную роль для управляющих активами, трейдеров и лиц, принимающих решения, начинают играть так называемые альтернативные данные, выходящие за рамки обычного ценообразования на рынке ценных бумаг, фундаментальных показателей деятельности компании или макроэкономических индикаторов. Основным источником таких данных являются социальные сети и мессенджеры. Инвестиционные фирмы нанимают экспертов по машинному обучению в области обработки естественного языка (*natural language processing*) для работы с финансовыми новостями, неструктурированными документами, отчетами по форме 10-K и пр. Такие крупные поставщики данных, как *Bloomberg*, *Thomson Reuters*, *RavenPack*, предоставляют обработанные данные о настроениях участников рынков в отношении новостей, адаптированные для моделей систематической торговли.

Следует отметить, что новые альтернативные наборы данных имеют определенные свойства. Многие из них неструктурированы, содержат нечисловые и (или) некатегориальные данные, например новостные статьи, голосовые записи или спутниковые изображения. В большинстве случаев это многомерные наборы данных (транзакции по кредитным картам), и количество переменных в них может значительно превышать количество наблюдений. Они также могут неявно содержать информацию о сетях агентов. Методы классической эконометрики не работают с такими наборами данных, так как часто они основаны на линейной алгебре и выдают ошибку, если количество переменных превышает количество наблюдений. Ковариационные матрицы не могут отразить топологические отношения, характеризующие сети. В свою очередь, методы машинного обучения предлагают числовую мощность и функциональную гибкость для выявления сложных закономерностей в многомерном пространстве. Последние достижения в области машинного обучения делают его применимым для оценки достоверности научных теорий, определения информационных переменных в объяснительных и (или) прогнозных целях, причинного вывода и визуализации больших, многомерных и комплексных наборов данных [19].

Моделирование ценообразования активов развивается в эмпирическом направлении. Используются обширные наборы характеристик компаний и сотни факторов для описания и понимания различий в ожидаемой доходности активов и моделирования динамики премии за риск инвестирования [20]. В сущности, измерение премии за риск — это проблема прогнозирования, так как премия за риск — это условное ожидание будущей реализованной избыточной прибыли. Высоко ценятся методологии, которые могут надежно соотнести избыточную доходность с аномалиями торговли. Машинное обучение обеспечивает нелинейный эмпирический подход к моделированию реальной доходности в зависимости от характеристик компаний. Например, М. Диксон и Н. Полсон [21] рассматривают формулировку моделей ценообразования активов для измерения премий за риск, используя 3290 активов из индекса Russell 1000 за период с декабря 1989 по январь 2018 г., и вводят нейронные сети в канонические системы ценообразования активов. Известно, что 49-факторная модель генерирует информационные коэффициенты в 1,5 раза лучше, чем обычная факторная модель на основе метода наименьших квадратов.

Появление индустрии финансовых технологий также обязано росту объемов данных и внедрению методов машинного обучения в технологичные бизнес-модели, цифровым инновациям в финансовом

секторе. Безусловно, центральное место в финансовых технологиях сейчас занимают криптовалюты и блокчейн-технологии, новые цифровые консультационные и торговые системы, мобильные платежные системы, одноранговое кредитование и краудфандинг. Критическим аспектом дизайна продукта и управления рисками, необходимого для бизнес-моделей, ориентированных на потребителя, является прогнозирование поведения агентов. Всем участникам рынка предоставляется четко определенный вариатив, но они имеют неизвестные экономические потребности и ограничения, а также поступают экономически нерационально, что затрудняет расчеты на основе классических моделей. Появление роботов-консультантов, предоставляющих финансовые консультации или услуги по управлению инвестиционным портфелем с минимальным вмешательством человека, позволило упростить многие процессы и расширить круг финансовых возможностей индивида. Основная масса таких технических решений предоставляет услуги по управлению инвестиционными портфелями, но есть и управляющие персональными финансами и пенсионными накоплениями (*Betterment, Wealthfront, WiseBanyan, FutureAdvisor, Bloom, Motif Investing, Personal Capital*). Уровень сложности технических решений по управлению инвестиционными портфелями и использования систем машинного обучения постоянно растет.

В 2018–2020 гг. около 2500 респондентов стали жертвами экономических преступлений, компании в среднем пережили по шесть инцидентов¹. Согласно исследованию к основным видам экономических преступлений относятся клиентское мошенничество (35 %), киберпреступность (34 %), незаконное присвоение активов (31 %), взяточничество и коррупция (30 %). Выявление экономических преступлений является одним из самых успешных приложений машинного обучения в сфере финансовых услуг. Постоянный рост объемов электронной торговли (в том числе из-за пандемии) ведет к новым видам финансового мошенничества и манипулирования рынком. Биржи, например, исследуют возможность использования глубокого обучения для противодействия спуфингу (*spoofing* – подмена) – ситуациям, когда человек или программа маскируются под другую программу путем фальсификации данных, что позволяет получить незаконные выгоды. Применение алгоритмов машинного обучения к опубликованным корпоративным данным позволяет выявлять закономерности, указывающие на мошенничество с ценными бумагами, манипуляции с раскрытием определенных статей финансовой отчетности, отражающих добросовестность компаний.

Блокчейн-технология (распределенный публичный реестр, регистрирующий транзакции) обеспечивает безопасную одноранговую связь между информационными блоками с меткой времени и данными транзакции. Первой децентрализованной цифровой валютой стал биткойн, использующий блокчейн для распределенного открытого хранения сведений о транзакциях, чтобы уменьшить недостатки финансовой индустрии. Современное представление данных допускает нетрадиционную форму финансовой эконометрики с упором на топологические сетевые структуры, а не только на ковариацию исторических временных рядов цен. Задачи исследований – изучение роли пользователей, организаций и их взаимодействия в формировании и динамике риска криптовалютных инвестиций, финансовая предиктивная аналитика.

С каждым годом финансы все больше полагаются на вычислительные методы. В то же время рост количества машиночитаемых данных для мониторинга, записи информации о действиях в финансовой системе и обмена ею имеет важные последствия для подхода к моделированию. Успех искусственного интеллекта и компьютерных алгоритмов обучения зависит от ряда факторов, выходящих за рамки компьютерного оборудования и программного обеспечения. Машины могут моделировать сложные многомерные процессы генерации данных, разворачивать миллионы модельных конфигураций, проводить робастные вычисления и корректировать модели в соответствии с новой информацией, при этом можно одновременно рассматривать несколько конкурирующих моделей для выбора оптимальной в конкретной рыночной ситуации. Вместе с тем внедрение машинного обучения постепенно меняет поведение участников рынка. Это позволяет рассуждать, экспериментировать и формировать свои взгляды на основе данных, эмпирически управлять процессами принятия решений.

Машинное обучение с учителем обычно представляет собой алгоритмическую форму статистической оценки модели, в которой процесс генерации данных полагается неизвестным. Выбор модели и вывод данных из нее автоматизированы с акцентом на обработку больших объемов данных. Это можно рассматривать как высокоэффективный метод сжатия данных для обеспечения предикторов в сложных условиях, когда отношения между входными и выходными переменными нелинейные, а пространство входов многомерное. Машинное обучение балансирует фильтрацию данных для принятия точных и надежных решений. Это в корне отличается от оценки методом максимального правдоподобия, который используется в стандартных статистических моделях, предполагающих, что данные были созданы моделью и обычно имеют трудности с подгонкой, особенно с многомерными наборами данных. Из-за сложности

¹ Global economic crime and fraud survey 2022. PwC [Electronic resource]. URL: <https://www.pwc.com/gx/en/services/forensics/economic-crime-survey.html> (date of access: 10.02.2022).

современных наборов данных, будь то списки заявок или многомерные финансовые временные ряды, делать вывод на основе известного процесса генерации данных часто невозможно. Даже если можно дать экономическую интерпретацию процесса генерации данных, его точная форма не может быть известна в каждый момент времени.

Парадигма машинного обучения для анализа данных сильно отличается от традиционной структуры статистического моделирования и тестирования. Такие традиционные метрики соответствия, как R^2 , t -значения, p -значения и статистическая значимость, заменяются прогнозированием вне выборки и пониманием компромисса между смещением и вариацией. Машинное обучение фокусируется на поиске структуры в больших наборах данных, а в качестве основных инструментов для выбора предиктора выступают регуляризация и отсев.

В табл. 1 представлено обобщенное сравнение модельного вывода, основанного на оценке методом максимального правдоподобия, и машинного обучения с учителем. В реальности эти подходы являются противоположными границами множества методов моделирования данных. Например, некоторые модели линейной регрессии, гребневая регрессия и их гибриды находятся где-то посередине. Они обеспечивают комбинацию объяснительной мощности метода оценки максимального правдоподобия при сохранении прогностической вневыборочной эффективности для многомерных наборов данных.

Статистические методы машинного обучения можно охарактеризовать как параметрические и непараметрические. К параметрическим методам относятся метод наименьших квадратов, полиномиальная регрессия, нейронные сети и скрытые марковские модели, а к непараметрическим – метод опорных векторов, гауссовы процессы (ядерные методы). Сравнение параметрических и непараметрических моделей представлено в табл. 2. Стоит отметить, что в зависимости от их настройки нейронные сети могут быть как параметрическими, так и непараметрическими.

С точки зрения парадигмы моделирования можно провести различие между вероятностными и детерминированными моделями. Первая группа рассматривает параметры как случайные, а вторая предполагает, что параметры заданы. В вероятностном моделировании особую нишу занимают так называемые модели пространства состояний, эволюция некоего ненаблюдаемого латентного процесса в которых приводит в движение определенный наблюдаемый процесс. Эволюция скрытого процесса и зависимость от него наблюдаемого процесса могут быть представлены в стохастических терминах, что помещает модели пространства состояний в область вероятностного моделирования. В свою очередь, детерминированная модель может давать вероятностные выходные данные, например, логистическая регрессия дает вероятность того, что ответ будет положительным с учетом входных переменных.

Основная проблема машинного обучения (особенно глубокого) – склонность к чрезмерной подгонке с учетом количества параметров в модели. В частотной статистике чрезмерная подгонка решается путем введения штрафа для функции правдоподобия. Распространенным подходом является выбор моделей на основе информационных критериев Акаике [22], предполагающих, что ошибка модели является гауссовой. Такие методы машинного обучения, как модель линейной регрессии LASSO (оператор наименьшей абсолютной усадки и выбора) и гребневая регрессия, более удобны для прямой оптимизации функции потерь со штрафом. Более того, подход не ограничивается допущениями о распределении ошибок моделирования. Регуляризация LASSO способствует более разреженной параметризации, а гребневая регрессия снижает значимость параметров. Регуляризация, возможно, является причиной успешности методов машинного обучения в сфере финансов. И наоборот, из-за ее отсутствия нейронные сети потеряли популярность в финансовой индустрии в 1990-х гг.

Сфера алгоритмической торговли – естественная площадка для машинного обучения. Торговые решения должны основываться на данных, а не на интуиции, следовательно, должна быть возможность автоматизировать этот процесс принятия решений с использованием алгоритма. Преимущества алгоритмической торговли включают в себя распознавание сложных рыночных моделей, уменьшение количества ошибок, производимых человеком, способность тестирования на материале исторических данных и т. д. В последнее время, когда все больше информации оцифровывается, возможности алгоритмической торговли резко увеличиваются.

Подводя итоги, следует отметить некоторые ключевые элементы машинного обучения. Машинное обучение с учителем представляет собой алгоритмический подход к статистическому выводу, который не зависит от процесса генерации данных; оценивает параметризованную карту между входами и выходами с учетом ее функциональной формы, определенной методологией; автоматизирует выбор модели, использует методы регуляризации и усреднения для перебора возможных моделей и получения модели с наилучшими характеристиками вне зависимости от выборки, а также хорошо подходит для больших многомерных данных.

Таблица 1

Сравнение статистического вывода на основе оценки методом максимального правдоподобия и машинного обучения с учителем

Table 1

Comparing statistical inference based on maximum likelihood estimation and machine learning with a teacher

Подход	Характеристика						
	Применяемые методы	Данные	База моделирования	Тип модели	Масштабируемость	Диагностика, проверка качества	Надежность
Статистический вывод	Вероятностные	Данные генерируются моделью	Информационный критерий	Обычно линейный	Ограничена данными более низкой размерности	Развернутая	Склонность к чрезмерной подгонке
Машинное обучение	Алгоритмические и вероятностные	Процесс создания данных принимается неизвестным	Численная оптимизация	Нелинейный	Масштабируется для входных данных большой размерности	Ограниченная	Работает вне зависимости от выборки
							Прогнозирование, часто с ограниченной объяснительной мощностью

Таблица 2

Сравнение параметрических и непараметрических моделей

Table 2

Comparison of parametric and nonparametric models

Модель	Характеристика		
	Параметры	Отклик	Гибкость
Параметрическая	Конечный набор параметров	Функция входных переменных и параметров	Не могут фиксировать сложные закономерности в больших данных
Непараметрические	Пространство параметров бесконечномерное	Не имеет жестких рамок	Структура обычно не определена априори и может усложняться с увеличением количества данных

За счет этого машинное обучение является более надежным подходом, чем многие методы параметрической финансовой эконометрики. Ключ к внедрению машинного обучения в финансах – возможность запускать машинное обучение вместе с параметрическими методами, наблюдая с течением времени различия и ограничения параметрического моделирования на основе метрик подгонки в выборке. Для характеристики данных и выбора алгоритма необходимо использовать статистические тесты. К другим преимуществам можно отнести способность легко расширять масштабы данных, но только в том случае, если данные имеют достаточно высокое качество и добавляют новый источник информации. В то же время применение машинного обучения требует серьезных навыков научного обоснования и не является панацеей для автоматического принятия решений.

Постепенные улучшения, конкурентные стратегии, сосредоточение внимания на ключевых компетенциях и снижение сложности и цены цепочки создания стоимости уже не являются достаточными ответами на снижение спроса компетентных клиентов компаний [23]. В большинстве секторов и отраслей произошел или происходит переход к цифровым технологиям с глубокими последствиями для их владельцев, руководителей и персонала. Благодаря лучшей подготовке к сбоям фирмы хотя бы могут получить дополнительное время для разработки стратегий преодоления неблагоприятных условий, подстроиться под сложившуюся ситуацию.

В ходе представленного исследования была выявлена тенденция к платформизации во многих сферах, поэтому ее следует рассматривать как основную концепцию и принимать в расчет, когда речь идет о цифровом разрушении отраслей. В этой области необходимы дальнейшие исследования, в частности определение отчетливых показателей для анализа и оценки того, какие отрасли с высокой вероятностью будут платформизированы в ближайшее время, идентификация роли различных структур в экосистеме конкретной платформы, спецификация цифровых бизнес-ситуаций.

Библиографические ссылки

1. Pecheranskyi I, Revenko A. Disruptive digital technologies as a means for destroying the foundations of oligarchomics: world experience and challenges for Ukraine. *Economic Annals-XXI*. 2019;9–10(179):31–39. DOI: 10.21003/ea.V179-03.
2. Khare A, Stewart B, Schatz R. *Phantom ex machina. Digital disruption's role in business model*. Cham: Springer; 2017. 354 p.
3. Coase RH. The nature of the firm. *Economica*. 1937;4(16):386–405.
4. Moore G. Cramming more components onto integrated circuits. *Electronics*. 1965;8(38):114–117. DOI: 10.1109/N-SSC.2006.4785860.
5. Hendler J, Golbeck J. Metcalfe's law, web 2.0, and the semantic web. *Journal of Web Semantics*. 2008;1(6):14–20. DOI: 10.1016/j.websem.2007.11.008.
6. Дудин МН, Шкодинский СВ, Усманов ДИ. Ключевые тенденции и закономерности развития цифровых бизнес-моделей банковских сервисов в индустрии 4.0. *Финансы: теория и практика*. 2021;5(25):59–78. DOI: 10.26794/2587-5671-2021-25-5-59-78.
7. Rochet J-C, Tirole J. Platform competition in two-sided markets. *Journal of the European Economic Association*. 2003;1(4):990–1029. DOI: 10.1162/154247603322493212.
8. Armstrong M. Competition in two-sided markets. *RAND Journal of Economics*. 2006;3(37):668–691.
9. Yoffie DB, Kwak M. With friends like these: the art of managing complementors. *Harvard Business Review*. 2006;9(84):88–98.
10. Hagiu A. Two-sided platforms: product variety and pricing structures. *Journal of Economics & Management Strategy*. 2009;4(18):1011–1043. DOI: 10.1111/j.1530-9134.2009.00236.x.
11. Galeotti A, Moraga-Gonzalez JL. Platform intermediation in a market for differentiated products. *European Economic Review*. 2009;4(53):417–428. DOI: 10.1016/j.euroecorev.2008.08.003.
12. Cennamo C, Santalo J. Platform competition: strategic trade-offs in platform markets. *Strategic Management Journal*. 2013;11(34):1331–1350. DOI: 10.1002/smj.2066.
13. Parker G, van Alstyne M. Innovation, openness, and platform control. *Management Science*. 2018;7(64):3015–3032. DOI: 10.1287/mnsc.2017.2757.
14. Jacobides MG, Cennamo C, Gawer A. Towards a theory of ecosystems. *Strategic Management Journal*. 2018;8(39):1–22. DOI: 10.1002/smj.2904.
15. Ceccagnoli M, Forman C, Huang P, Wu DJ. Co-creation of value in a platform ecosystem: the case of enterprise software. *MIS Quarterly*. 2012;1(36):263–290. DOI: 10.2307/41410417.
16. Gawer A, Cusumano M. How companies become platform leaders. *MIT Sloan Management Review*. 2008;2(49):28–35.
17. Wareham J, Fox PB, Cano Giner JL. Technology ecosystem governance. *Organization Science*. 2014;4(25):1195–1215. DOI: 10.2139/ssrn.2201688.
18. Flood M, Jagdish HV, Raschid L. Big data challenges and opportunities in financial stability monitoring. *Financial Stability Review*. 2016;20:129–142.
19. López de Prado MM. Beyond econometrics: a roadmap towards financial machine learning. *Econometric Modelling: Theoretical Issues in Microeconometrics eJournal* [Internet]. 2019 [cited 2022 February 10]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3365282. DOI: 10.2139/ssrn.3365282.
20. Gu S, Kelly BT, Xiu D. Empirical asset pricing via machine learning. *The Review of Financial Studies*. 2020;5(33):2223–2273.
21. Dixon MF, Polson NG. Short communication: deep fundamental factor models. *SIAM Journal on Financial Mathematics* [Internet]. 2020 [cited 2022 February 10];3(11). Available from: <https://epubs.siam.org/doi/abs/10.1137/20M1330518>. DOI: 10.1137/20M1330518.

22. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. *Selected Papers of Hirotugu Akaike. Springer Series in Statistics*. New York: Springer; 1998. p. 199–213.
23. Бондаренко ИА. Актуальные вопросы распространения и внедрения цифровых бизнес-моделей в России. *Естественно-гуманитарные исследования*. 2021;3(35):63–67. DOI: 10.24412/2309-4788-2021-11126.

References

1. Pecheranskyi I, Revenko A. Disruptive digital technologies as a means for destroying the foundations of oligarchomics: world experience and challenges for Ukraine. *Economic Annals-XXI*. 2019;9–10(179):31–39. DOI: 10.21003/ea.V179-03.
2. Khare A, Stewart B, Schatz R. *Phantom ex machina. Digital disruption's role in business model*. Cham: Springer; 2017. 354 p.
3. Coase RH. The nature of the firm. *Economica*. 1937;4(16):386–405.
4. Moore G. Cramming more components onto integrated circuits. *Electronics*. 1965;8(38):114–117. DOI: 10.1109/N-SSC.2006.4785860.
5. Hendler J, Golbeck J. Metcalfe's law, web 2.0, and the semantic web. *Journal of Web Semantics*. 2008;1(6):14–20. DOI: 10.1016/j.websem.2007.11.008.
6. Dudin MN, Shkodinskii SV, Usmanov DI. Key trends and regulations of the development of digital business models of banking services in Industry 4.0. *Finance: Theory and Practice*. 2021;5(25):59–78. Russian. DOI: 10.26794/2587-5671-2021-25-5-59-78.
7. Rochet J-C, Tirole J. Platform competition in two-sided markets. *Journal of the European Economic Association*. 2003;1(4):990–1029. DOI: 10.1162/154247603322493212.
8. Armstrong M. Competition in two-sided markets. *RAND Journal of Economics*. 2006;3(37):668–691.
9. Yoffie DB, Kwak M. With friends like these: the art of managing complementors. *Harvard Business Review*. 2006;9(84):88–98.
10. Hagiu A. Two-sided platforms: product variety and pricing structures. *Journal of Economics & Management Strategy*. 2009;4(18):1011–1043. DOI: 10.1111/j.1530-9134.2009.00236.x.
11. Galeotti A, Moraga-Gonzalez JL. Platform intermediation in a market for differentiated products. *European Economic Review*. 2009;4(53):417–428. DOI: 10.1016/j.eurocorev.2008.08.003.
12. Cennamo C, Santalo J. Platform competition: strategic trade-offs in platform markets. *Strategic Management Journal*. 2013;11(34):1331–1350. DOI: 10.1002/smj.2066.
13. Parker G, van Alstyne M. Innovation, openness, and platform control. *Management Science*. 2018;7(64):3015–3032. DOI: 10.1287/mnsc.2017.2757.
14. Jacobides MG, Cennamo C, Gawer A. Towards a theory of ecosystems. *Strategic Management Journal*. 2018;8(39):1–22. DOI: 10.1002/smj.2904.
15. Ceccagnoli M, Forman C, Huang P, Wu DJ. Co-creation of value in a platform ecosystem: the case of enterprise software. *MIS Quarterly*. 2012;1(36):263–290. DOI: 10.2307/41410417.
16. Gawer A, Cusumano M. How companies become platform leaders. *MIT Sloan Management Review*. 2008;2(49):28–35.
17. Wareham J, Fox PB, Cano Giner JL. Technology ecosystem governance. *Organization Science*. 2014;4(25):1195–1215. DOI: 10.2139/ssrn.2201688.
18. Flood M, Jagadish HV, Raschid L. Big data challenges and opportunities in financial stability monitoring. *Financial Stability Review*. 2016;20:129–142.
19. López de Prado MM. Beyond econometrics: a roadmap towards financial machine learning. *Econometric Modelling: Theoretical Issues in Microeconometrics eJournal* [Internet]. 2019 [cited 2022 February 10]. Available from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3365282. DOI: 10.2139/ssrn.3365282.
20. Gu S, Kelly BT, Xiu D. Empirical asset pricing via machine learning. *The Review of Financial Studies*. 2020;5(33):2223–2273.
21. Dixon MF, Polson NG. Short communication: deep fundamental factor models. *SIAM Journal on Financial Mathematics* [Internet]. 2020 [cited 2022 February 10];3(11). Available from: <https://epubs.siam.org/doi/abs/10.1137/20M1330518>. DOI: 10.1137/20M1330518.
22. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. *Selected Papers of Hirotugu Akaike. Springer Series in Statistics*. New York: Springer; 1998. p. 199–213.
23. Bondarenko IA. Actual issues of distribution and implementation of digital business models in Russian Federation. *Natural Humanitarian Studies*. 2021;3(35):63–67. Russian. DOI: 10.24412/2309-4788-2021-11126.

Статья поступила в редколлегию 13.02.2022.
Received by editorial board 13.02.2022.