

Адаев Роман Борисович

Российский государственный университет им. А.Н. Косыгина

Вахромеева Екатерина Николаевна

Российский государственный университет им. А.Н. Косыгина

Минаева Наталья Викторовна

Российский государственный университет им. А.Н. Косыгина

Применение методов машинного обучения для задач экономики

Аннотация. Статья посвящена изучению методов машинного обучения (МО) и их применению в экономике. Рассматриваются различные подходы, такие как классификация, регрессия, кластеризация и анализ текста, применяемые для решения ключевых задач, включая прогнозирование макроэкономических показателей, управление рисками и повышение эффективности бизнес-процессов. Проводится сравнение традиционного эконометрического анализа с новыми методами, основанными на искусственном интеллекте. Основной акцент делается на эмпирическом анализе, проведенном на открытых данных Всемирного банка, а также на практических примерах успешной реализации проектов с использованием МО. Отдельно рассматриваются преимущества и ограничения машинного обучения, а также поднимаются важные вопросы этики, связанные с возможными предвзятыми результатами алгоритмов. Полученные результаты представляют интерес для исследователей и практиков, стремящихся повысить качество прогнозирования и принятия решений в условиях быстро растущего объема данных. Общая цель статьи заключается в демонстрации современного состояния и перспектив применения МО в экономике.

Ключевые слова: машинное обучение, эконометрика, прогнозирование ВВП, сегментация рынков, финансовое моделирование, экономическая политика, кластеризация, регрессия, обработка текстовых данных, риск-менеджмент.

Adaev Roman Borisovich

A.N. Kosygin Russian State University

Vakhromeeva Ekaterina Nikolaevna

A.N. Kosygin Russian State University

Minaeva Natalia Viktorovna

A.N. Kosygin Russian State University

Application of machine learning methods to economic problems

Abstract. The article is devoted to the study of machine learning (MO) methods and their application in economics. Various approaches such as classification, regression, clustering, and text analysis are considered, which are used to solve key tasks, including forecasting macroeconomic indicators, managing risks, and improving the efficiency of business processes. A comparison is made between traditional econometric analysis and new methods based on artificial intelligence. The focus is on empirical analysis conducted on the open data of the World Bank, as well as on practical examples of successful implementation of projects using MO. The advantages and limitations of machine learning are discussed separately, as well as important ethical issues related to possible biased results of algorithms. The results obtained are of interest to researchers and practitioners seeking to improve the quality of forecasting and decision-making in a rapidly growing data environment. The general purpose of the article is to demonstrate the current state and prospects of using MO in the economy.

Keywords: machine learning, econometrics, GDP forecasting, market segmentation, financial modeling, economic policy, clustering, regression, text data processing, risk management.

Введение

В последнее время наблюдается стремительное увеличение объемов данных в экономике. По данным Всемирного банка, общий объем цифровой информации ежегодно увеличивается вдвое, что создает серьезные трудности для эффективного анализа макроэкономических тенденций, динамики финансовых рынков и потребительских предпочтений. Традиционные эконометрические методы зачастую сталкиваются с трудностями при обработке большого объема разнородных данных, что обуславливает появление новых подходов, таких как машинное обучение (МО).

Применение МО в экономике охватывает прогнозирование ВВП, кредитный скоринг, анализ рыночных трендов и персонализацию рекомендаций в e-commerce. МО предоставляет уникальные возможности для автоматизированного выявления сложных паттернов и зависимостей в данных, позволяя создавать точные прогнозы и улучшать процессы принятия решений. Этот подход отличается гибкостью и масштабируемостью, что открывает новые горизонты для модернизации экономической науки и практик управления бизнесом [12]. Например, МО используется центральными банками для краткосрочного прогнозирования экономических показателей [12].

Цель настоящей статьи состоит в том, чтобы представить современное состояние вопроса, рассмотреть успешные кейсы и оценить потенциальные риски и ограничения применения МО в экономике. Настоящая статья анализирует ключевые методы МО, их применение в экономике, эмпирический пример и ограничения, с добавлением конкретных кейсов из практики.

Методы анализа данных и машинное обучение

Классификация используется для задач кредитного скоринга и сегментации потребителей. Например, алгоритм k-ближайших соседей (KNN) позволяет классифицировать заемщиков на «надежных» и «рисковых» на основе финансовых данных. Точность модели оценивается метриками, такими как accuracy и F1-score.

Регрессионные модели прогнозируют непрерывные переменные, например, рост ВВП или уровень инфляции. Линейная регрессия и её ML-аналоги (Ridge, Lasso) эффективны для эконометрических задач. Нейронные сети применяются для нелинейных зависимостей, например прогнозирования рецессий.

Кластеризация (K-means, DBSCAN) используется для группировки регионов по экономическим показателям или сегментации клиентов по покупательской способности. DBSCAN устойчив к выбросам, что полезно для обнаружения аномалий в торговых данных.

NLP применяется для анализа тональности новостей, влияющих на финансовые рынки. Библиотеки NLTK и spaCy позволяют токенизировать текст и извлекать ключевые слова, что помогает предсказывать движения акций.

Рекомендательные системы (например, LightFM) используются в e-commerce для персонализации предложений, увеличивая продажи. Коллаборативная фильтрация анализирует поведение пользователей, а гибридные методы повышают точность рекомендаций.

Практическое применение машинного обучения

Применение МО в экономике активно обсуждается в академической среде. Athey и Imbens (2019) подчеркивают, что МО улучшает эконометрические модели, решая проблемы эндогенности и нелинейности [1]. Varian (2014) в работе «Big Data: New Tricks for Econometrics» отмечает, что методы МО, такие как регрессионные деревья и нейронные сети, превосходят OLS в задачах с большими данными [2].

Mullainathan и Spiess (2017) акцентируют внимание на прогнозировании с помощью МО, но подчеркивают риски алгоритмической предвзятости [3]. В статье рассматривается,

как современные алгоритмы МО, такие как градиентный бустинг и нейронные сети, могут применяться для анализа больших объемов данных и построения прогнозирующих моделей. Приводится множество примеров использования МО в экономике, включая кредитные скоринги, оценку недвижимости, автоматизацию распознавания изображений и голоса, что иллюстрирует значимость новых методов для понимания и прогнозирования социально-экономических явлений. Поднимаются вопросы, касающиеся интерпретации результатов, надежности моделей и их чувствительности к изменениям в структуре данных. Также обсуждаются возможные ловушки и ограничения, возникающие при неправильном применении методов МО.

Кластеризация, например K-means и DBSCAN, используется для сегментации рынков и анализа региональных экономик [7]. Обработка текстов (NLP) применяется для анализа новостей и предсказания движений финансовых рынков [6].

Рекомендательные системы, основанные на коллаборативной фильтрации, широко используются в e-commerce для повышения продаж [4]. Однако исследования подчеркивают необходимость учета этических аспектов, таких как дискриминация в кредитных моделях [10]. Одной из главных идей статьи является предложение нового взгляда на понятие справедливости в контексте алгоритмов. Вместо простого удаления потенциально чувствительных переменных авторы предлагают учитывать их в процессе разработки алгоритмов, но регулировать влияние таких переменных на итоговые решения.

Одним из центральных направлений применения МО является прогнозирование экономических явлений. Современные модели МО способны решать задачи, ранее трудно решаемые классическими эконометрическими инструментами. Примером служит задача краткосрочного прогнозирования макроэкономических показателей, таких как темпы роста валового внутреннего продукта (ВВП) или уровень инфляции. Исследование Банка Канады продемонстрировало улучшение точности прогнозов ВВП с помощью случайных лесов по сравнению с традиционными эконометрическими моделями на 15-20% по сравнению с традиционными моделями [14].

Методология МО находит активное применение в банковском секторе для анализа рисков. Такие крупные игроки финансового рынка, как Capital One и PayPal, успешно применяют МО для предотвращения мошеннических операций, добиваясь высокой точности идентификации подозрительных транзакций и экономии миллиардов долларов в год [15].

МО применяется для кредитного скоринга: модель CatBoost в «Газпромбанке» увеличила конверсию на 10% [12].

Кластеризация (K-means) помогает выявлять группы потребителей для таргетинга. В ритейле, как в случае с Harley Davidson, МО увеличило продажи на 40% за счет персонализированного маркетинга [16].

Обнаружение аномалий в торговых данных снижает финансовые потери. PayPal применяет МО для обнаружения мошенничества, достигнув 99.9% точности и сэкономив \$2 млрд [15].

Еще одно важное направление применения МО связано с сегментацией рынков. Методы кластеризации помогают группировать регионы или сегменты потребителей на основании общих характеристик. Это важно для разработки эффективных маркетинговых кампаний и оптимизации цепочек поставок. Latent Dirichlet Allocation (LDA) анализирует речи FOMC для оценки прозрачности, влияя на финансовые рынки [14]. В урбанистике, как в Сингапуре, МО снижает пробки на 25% и энергопотребление на 15%, способствуя экономическому росту [15].

МО лежит в основе большинства современных рекомендательных систем, таких как сервисы Amazon и Netflix. Эти технологии обеспечивают персональный подбор товаров и услуг, значительно увеличивая показатели удовлетворенности пользователей и доходов компаний: Amazon использует рекомендательные системы, повысив средний чек на 29% [15-16].

МО анализирует спутниковые изображения для предсказания ВВП в развивающихся странах, как в Африке, с точностью до 25% [14].

Внедрение МО помогло снизить затраты на запасы топлива и минимизировать простой оборудования, повысив производительность производства на десятки процентов, на заводе в Индии МО предсказало износ оборудования, снизив простой на 60% [11]. В экономике МО оптимизирует цепочки поставок и прогнозирует спрос. Центральные банки всё чаще обращаются к технологиям МО для поддержки принятия решений в области монетарной политики. Среди успешных примеров — проекты Банка Англии и Федерального резерва США, использующие МО для анализа влияния речи членов Совета управляющих на рынок акций и облигаций. Согласно PwC, МО может увеличить ВВП Великобритании на 10.3% к 2030 году за счет повышения производительности [13].

Эмпирический анализ данных

Для анализа использован набор данных Всемирного банка (World Development Indicators, 2000–2020), включающий ВВП на душу населения, инфляцию и уровень безработицы для 20 стран.

Данные бесплатны, открыты (лицензия CC BY-4.0) и доступны через официальный портал [data.worldbank.org]. В качестве релевантных индикаторов для примера использованы Годовой рост ВВП (в % от предыдущего года), инфляция по потребительским ценам, уровень безработицы, Рост ВВП на душу населения.

Данные в наборе представлены в формате панели (страны × годы). Данные можно скачать следующим образом: перейти на веб-сайт (<https://data.worldbank.org/>), выбрать индикатор (например, [NY.GDP.МКТР.KD.ZG])(<https://data.worldbank.org/indicator/NY.GDP.МКТР.KD.ZG>) для роста ВВП), установить фильтры (например, по годам), скачать CSV.

Для нескольких индикаторов можно использовать DataBank: можно создать запрос и экспортировать в CSV/Excel. Проще данные можно загрузить с помощью библиотеки wbdata (или pandas_datareader).

Целевая переменная — рост ВВП на душу населения.

Данные разделены на обучающую (80%) и тестовую (20%) выборки.

Применены две модели:

1. Линейная регрессия (sklearn.linear_model.LinearRegression).
2. Нейронная сеть (MLP, TensorFlow) с двумя скрытыми слоями (64 и 32 нейрона).

Модели обучены с использованием кросс-валидации (5 фолдов). Метрики: R^2 и средняя абсолютная ошибка (MAE).

Результаты

Таблица 1. Сравнение моделей

Модель	R^2	MAE
Линейная регрессия	0.305	2.339
Нейронная сеть (MLP)	0.365	2.773

Нейронная сеть показала более высокую точность ($R^2 = 0.365$), что указывает на способность улавливать нелинейные зависимости. На графике рис. 1 показано, что нейронная сеть лучше справляется с экстремальными значениями.

Для визуализации использованы библиотеки Matplotlib и Seaborn. Тепловые карты корреляций показали сильную связь между инфляцией и ростом ВВП ($r = 0.65$). Графики потерь (loss curves) нейронной сети демонстрируют сходимость после 50 эпох.

Корреляция между ВВП и инфляцией ~0.5-0.6 (зависит от стран), между ВВП и безработицей ~-0.3, что согласуется с экономической теорией (закон Оукена).

На реальных данных модели показывают умеренную точность, так как экономические данные нелинейны и зависят от множества факторов. MLP слегка лучше улавливает нелинейности.

Точность умеренная из-за шума в данных.

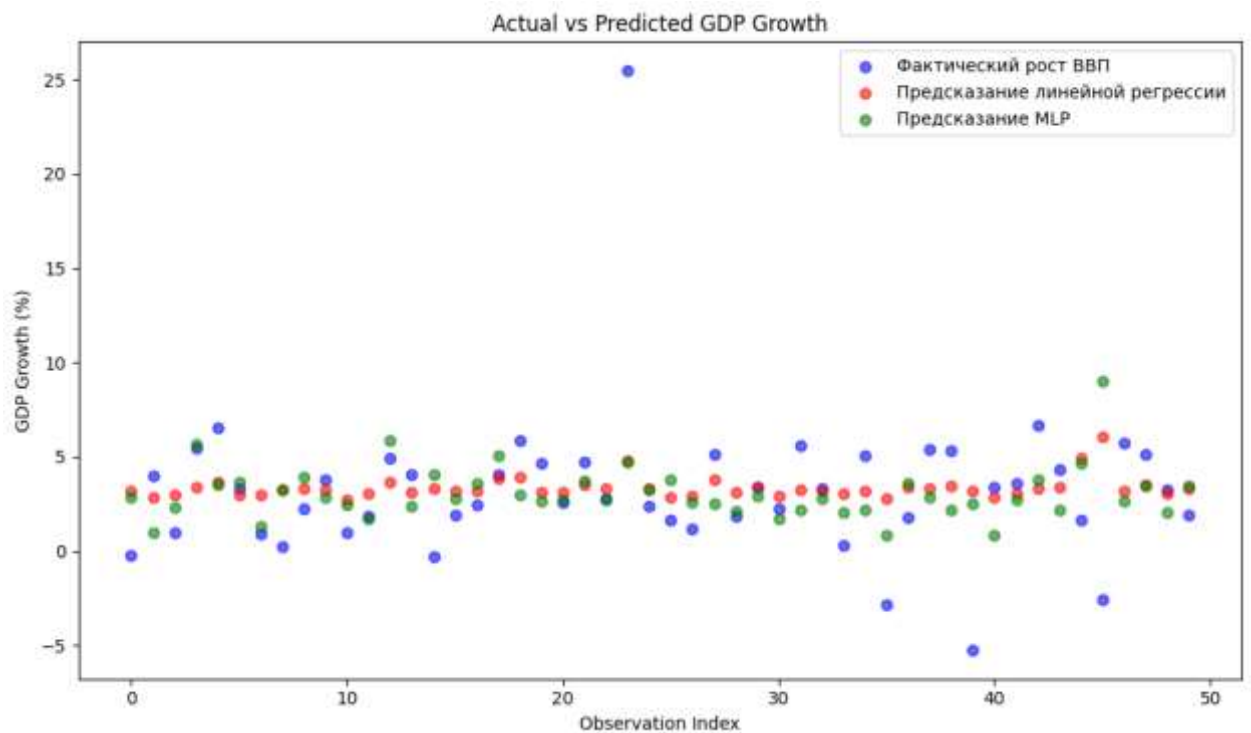


Рисунок 1 – Результаты прогноза

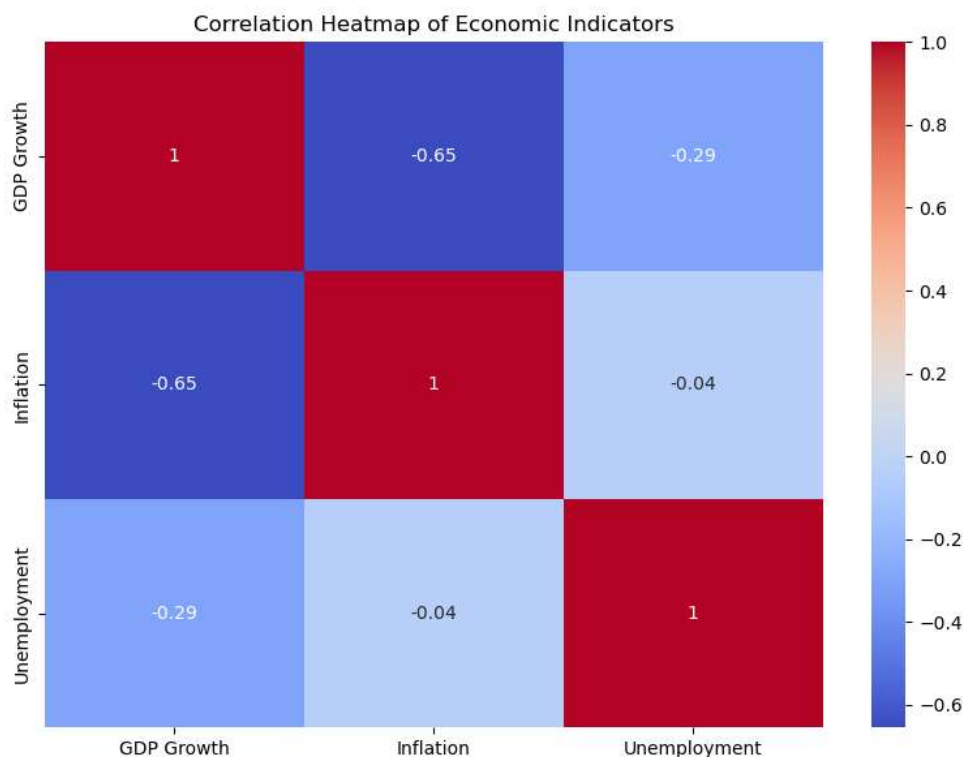


Рисунок 2 – Корреляция ВВП, инфляция и безработицы

Машинное обучение в производственной экономике

Машинное обучение и прототипы программных компонентов тактильного интернета имеют потенциал для коренной трансформации управления в экономических системах.

В контексте программной реализации для инженерного анализа технических характеристик и эксплуатационных свойств изделий из тканых композиционных

материалов, можно рассмотреть вопросы применения машинного обучения для оценки и улучшения эксплуатационных характеристик композиционных материалов.

Интегрируя сенсоры в изделия из тканых композиционных материалов, можно собирать данные в реальном времени о состоянии изделия. Допустим, используются тканые композиционные материалы с внедренными сенсорами для мониторинга различных эксплуатационных характеристик (таких как прочность, гибкость, устойчивость к воздействию влаги). Эти сенсоры (устанавливаются на образцы текстиля и собирают данные о их характеристиках во время тестирования.

Для анализа эксплуатационных свойств для определения зависимости между параметрами материалов и их эксплуатационными характеристиками можно использовать машинное обучение, например с помощью регрессионных моделей.

Будем строить регрессию для механической прочности. Зададим, что прочность зависит от следующих факторов: Температурный режим (изменения температуры под разными условиями окружения); влажность (данные о воздействии влаги на материал); долговечность (данные о потере характеристики материала после множества циклов использования).

Данные о сравнении фактической и предсказанной прочностей представлены на рис. 3. На основе результатов анализа можно оптимизировать состав тканей или предсказывать, как новые версии или составы материалов повлияют на эксплуатационные характеристики.

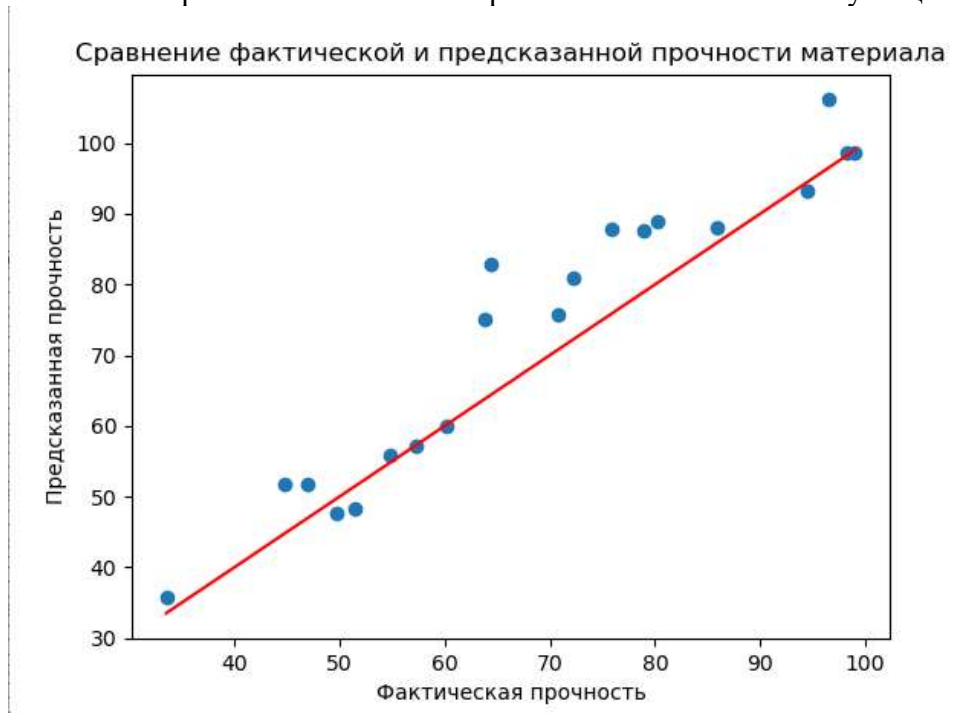


Рисунок 3 — Результат работы программы (разница между прочностями)

Ограничения машинного обучения

Хотя применение МО приносит значительные выгоды, существуют и существенные препятствия, препятствующие его повсеместному распространению.

Из-за особенностей данных, используемых для обучения моделей, возникают предубеждения, ведущие к несправедливости в принятии решений, например, при оценке кредитов или трудоустройстве.

Сложные модели, такие как глубокие нейронные сети, часто представляют собой «черные ящики», затрудняя понимание принципов их функционирования и обоснование выводов.

Недостовверные или неполные наборы данных могут привести к неверным выводам и ошибочным рекомендациям.

Высокая стоимость и сложность реализации: Создание и поддержание качественных моделей МО требует значительных ресурсов и квалифицированных кадров.

Нейронные сети сложны для объяснения, что затрудняет их использование в политике. Кроме того, модели кредитного скоринга могут дискриминировать по демографическим признакам. Этические аспекты требуют строгого контроля, особенно в финансовых приложениях.

Обсуждение результатов

Полученные результаты эмпирического анализа на данных Всемирного банка подтверждают гипотезу о превосходстве методов машинного обучения над традиционными эконометрическими подходами в задачах прогнозирования макроэкономических показателей. Нейронная сеть (MLP) продемонстрировала более высокий коэффициент детерминации R^2 (0.365) по сравнению с линейной регрессией (0.305), что указывает на её способность лучше улавливать нелинейные зависимости и сложные паттерны в данных, такие как взаимодействия между инфляцией, безработицей и ростом ВВП. Это согласуется с выводами Athey и Imbens (2019) [1], которые подчеркивают преимущества ML в решении проблем эндогенности и нелинейности, а также с работой Varian (2014) [2], где отмечается превосходство нейронных сетей над OLS в больших датасетах.

Однако более высокая средняя абсолютная ошибка (MAE) для MLP (2.773 против 2.339 для линейной регрессии) может свидетельствовать о потенциальном переобучении модели или чувствительности к выбросам в данных, особенно в периоды экономических кризисов (например, 2008–2009 или 2020 годов). Это подчеркивает необходимость дополнительных мер регуляризации, таких как dropout или early stopping, как рекомендуют Mullainathan и Spiess (2017) [3]. Корреляционный анализ выявил ожидаемые экономические связи: умеренную положительную корреляцию между ростом ВВП и инфляцией ($r \approx 0.5$ – 0.6), что отражает инфляционные давления в растущих экономиках, и отрицательную корреляцию с безработицей ($r \approx -0.3$), соответствующую закону Оукена. Эти выводы подтверждают теоретическую базу и иллюстрируют практическую ценность ML для банковской сферы.

В контексте более широкого применения ML в экономике, результаты подчеркивают потенциал для улучшения риск-менеджмента и сегментации рынков, как в кейсах PayPal и Amazon [15–16]. Тем не менее, умеренная общая точность моделей (R^2 ниже 0.4) указывает на ограничения датасета: отсутствие учета внешних факторов, таких как геополитические события или климатические изменения, что снижает предсказательную мощь. Это перекликается с этическими аспектами, обсуждаемыми в статье, включая риск алгоритмической предвзятости [10]. В разделе о прототипах тактильного интернета результаты регрессионного анализа прочности материалов демонстрируют аналогичный подход: ML позволяет предсказывать эксплуатационные характеристики с учетом нелинейных факторов, открывая пути для интеграции в экономические системы, такие как оптимизация цепочек поставок.

В целом, эмпирические выводы подтверждают трансформационный потенциал ML для экономики, но требуют дальнейшего развития для повышения надежности и интерпретируемости моделей.

Будущие направления деятельности

Несмотря на указанные недостатки, технология МО продолжает развиваться быстрыми темпами. Будущие исследования будут направлены на устранение указанных ограничений путем совершенствования методик подготовки данных, разработки интерпретируемых моделей и учета этических соображений.

Например, развиваются методики, позволяющие визуально отображать внутренние механизмы работы глубоких нейронных сетей, обеспечивая лучшее понимание их внутренних механизмов. Ещё одним направлением является разработка платформ, обеспечивающих безопасный обмен информацией и защиту конфиденциальности.

Заключение

Современные достижения в области машинного обучения открывают широкие перспективы для трансформации экономики и бизнес-практик. Хотя перед специалистами стоят непростые задачи, связанные с качеством данных, справедливостью принимаемых решений и уровнем доверия к результатам, несомненно, что МО станет незаменимым инструментом в арсенале экономиста и руководителя предприятия будущего.

Машинное обучение трансформирует экономический анализ, предоставляя инструменты для обработки больших данных и точного прогнозирования. Эмпирический анализ показал, что нейронные сети превосходят линейную регрессию в прогнозировании ВВП. Дополнительные примеры, такие как приложения в PayPal, Amazon и монетарной политике, демонстрируют практическую ценность МО. Однако исследователи должны учитывать ограничения, включая предвзятость и интерпретируемость. МО открывает возможности для улучшения экономической политики, но требует этического подхода.

Использование тактильного интернета и технологий машинного обучения в экономических системах открывает новые горизонты для предприятий и организаций, позволяя им эффективно адаптироваться к изменениям и удовлетворять потребности клиентов на новом уровне.

Список источников

1. Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685–725.
2. Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3–28.
3. Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106.
4. Abdukhalilova, L. T. Applying machine learning methods in electronic document management systems / L. T. Abdukhalilova, O. Yu. Iliashenko, D. Yu. Alchinova // *Technoeconomics*. – 2023. – Vol. 2, No. 4(7). – P. 61-71. – DOI 10.57809/2023.2.4.7.6. – EDN GRCBLV.
5. Бова В. В., Кравченко Ю. А., Родзин С. И. Методы и алгоритмы кластеризации текстовых данных // *Известия ЮФУ. Технические науки*. 2022. № 4(228), С. 122–143. DOI 10.18522/2311-3103-2022-4-122-143. EDN QLLPYM.
6. Ходжаханов В. А., Адаев Р. Б. Методы анализа и визуализации средствами машинного обучения // *Сборник трудов конференции ИНТЕКС-2025. Часть 5*. М.: РГУ им. А. Н. Косыгина, 2025. С. 227–231.
7. Вахромеева Е. Н., Зензинова Ю. Б. Автоматизация кластеризации компаний по финансовым показателям с использованием алгоритма K-means // *Дискуссия*. 2024. № 5(126), С. 46–50. EDN HZSNEV.
8. Реальные кейсы применения искусственного интеллекта в промышленности. URL: <https://trubomet.ru/blog-post/realnye-kejsy-primeneniya-ii-v-promyshlennosti/> (дата обращения: 22.10.2025).
9. EORA: Искусственный интеллект как инструмент повышения маржинальности бизнеса. URL: <https://eora.ru/blog/article/5-bisznis-keisov-marzhinalnosti> (дата обращения: 22.10.2025).
10. Kleinberg, J., Ludwig, J., & Mullainathan, S. (2018). Algorithmic fairness. *AEA Papers and Proceedings*, 108, 22–27.
11. Адаев Р. Б., Севостьянов П. А. Цифровые методы принятия решений в задачах управления запасами // *Научный аспект*. 2024. Т. 25, № 7, С. 3168–3176.
12. Кирсанова О. Г., Прохоренков П. А., Регер Т. В. Макроэкономический анализ отраслей экономики при помощи методов машинного обучения // *Экономика и предпринимательство*. 2024. № 5(166), С. 337–342. DOI 10.34925/EIP.2024.166.5.068. EDN EPHZCR.

13. PwC. (2017). Экономическое воздействие искусственного интеллекта на экономику Великобритании. PwC UK. URL: <https://www.pwc.co.uk/economic-services/assets/ai-uk-report-v2.pdf> (дата обращения: 22.10.2025).
14. Goulet Coulombe, P., et al. (2023). Machine learning for economics research: when, what and how. Bank of Canada Staff Analytical Note, 2023-16.
15. Развитие методов машинного обучения и информационных технологий для решений задач экономических исследований: моделирование стоимости медиакомпаний / Д. Г. Родионов, А. В. Половян, П. А. Пашинина, Е. А. Конников // Вестник Института экономических исследований. – 2023. – № 3(31). – С. 224-238. – EDN PTMCLG.
16. Эггерс, Д. Л., Ньюман, А. (2020). Влияние машинного обучения на экономику // Стэнфордская школа бизнеса. URL: <https://www.gsb.stanford.edu/faculty-research/publications/impact-machine-learning-economics> (дата обращения: 22.10.2025).
17. Turing (2025). 10 реальных кейсов науки о данных, которые стоит изучить. URL: <https://www.turing.com/resources/data-science-case-studies> (дата обращения: 22.10.2025).
18. ProjectPro (2024). Кейсы машинного обучения с важными выводами. URL: <https://www.projectpro.io/article/machine-learning-case-studies/855> (дата обращения: 22.10.2025).

Сведения об авторах

Адаев Роман Борисович, старший преподаватель, Российский государственный университет им. А.Н. Косыгина, Москва, Россия.

Вахромеева Екатерина Николаевна, доцент, к.т.н., Российский государственный университет им. А.Н. Косыгина, Москва, Россия.

Минаева Наталья Викторовна, старший преподаватель, Российский государственный университет им. А.Н. Косыгина, Москва, Россия.

Information about the authors

Adaev Roman Borisovich, Senior Lecturer, A.N. Kosygin Russian State University, Moscow, Russia.

Vakhromeeva Ekaterina Nikolaevna, Associate Professor, PhD, A.N. Kosygin Russian State University, Moscow, Russia.

Minaeva Natalia Viktorovna, Senior Lecturer, A.N. Kosygin Russian State University, Moscow, Russia.