

УДК 621-039-542

<https://www.doi.org/10.47813/nto.5.2024.3002>

EDN [YHEYBQ](#)

Возможности использования искусственного интеллекта и машинного обучения для анализа фондового рынка

Юлия Сергеевна Варламова¹, Егор Владимирович Заведеев^{2*}

¹Сургутский государственный университет, Институт экономики и управления СурГУ, студент, Сургут, Россия

²Сургутский государственный университет, Институт экономики и управления СурГУ, доцент кафедры менеджмента и бизнеса, Сургут, Россия

*E-mail: zavedeev@mail.ru

Аннотация. Цель. Исследовать возможности применения ИИ (искусственного интеллекта) и машинного обучения при прогнозировании котировок акций, облигаций, фьючерсов и пр. финансовых инструментов и построения оптимальных инвестиционных стратегий на фондовых рынках. Проблема. Обычные, традиционные методы анализа фондового рынка действуют либо только в рамках технического анализа, либо только фундаментального и ограничены в возможностях учитывать действия экономических, политических и пр. событий при построении математических моделей для прогнозирования ситуации на фондовых рынках. Гипотеза. Применение ИИ и машинного обучения для анализа фондового рынка и прогнозирования трендов изменения цен финансовых инструментов, позволит улучшить точность прогнозирования, оптимизировать инвестиционные стратегии, а также поможет определить рискованные инвестиции.

Ключевые слова: искусственный интеллект, машинное обучение, фондовый рынок, анализ данных, алгоритмы машинного обучения.

The possibilities of using artificial intelligence and machine learning to analyze the stock market

Julia Sergeevna Varlamova¹, Egor Vladimirovich Zavedeev^{2*}

¹Surgut State University, Institute of Economics and Management, student, Surgut, Russia

²Surgut State University, Institute of Economics and Management, Associate Professor of the Department of Management and Business, Surgut, Russia

*E-mail: zavedeev@mail.ru

Abstract. Goal. To explore the possibilities of using AI (artificial intelligence) and machine learning in predicting stock prices, bonds, futures, etc. of financial instruments and building optimal investment strategies in stock markets. Problem. Conventional, traditional methods of stock market analysis operate either only within the framework of technical analysis, or only fundamental and are limited in their ability to take into account the actions of economic, political, etc. events when building mathematical models to predict the situation on stock markets. Hypothesis. The use of AI and machine learning to analyze the stock market and predict trends in the prices of financial instruments will improve the accuracy of forecasting, optimize investment strategies, and also help identify risky investments.

Keywords: artificial intelligence, machine learning, stock market, data analysis, machine learning algorithms.

1. Введение

Фондовый рынок – это сложная, постоянно подверженная изменениям под действием множества факторов система, от успешного прогнозирования состояния которой напрямую зависит эффективность инвестиционных стратегий, в соответствии с которыми работают участники фондовых рынков. К факторам, влияющим на состояние фондового рынка следует отнести: экономические и политические события, инвестиционный климат, уровень инфляции, ставку рефинансирования центрального банка, настроения инвесторов и целый ряд прочих факторов.

2. Материалы и методы

Традиционные методы прогнозирования состояния фондового рынка основаны на инструментах технического или фундаментального анализа. Первые предполагают построение линий трендов, расчет большого количества разнообразных трендовых индикаторов, сопоставление текущих продаж с продажами в прошлые периоды. Вторая группа методов предполагает анализ финансового состояния предприятий, расчет показателей доходности, определение прибыли на одну акцию и прогнозной величины потенциальных дивидендов.

Методы ИИ и машинное обучение при прогнозировании котировок позволяет объединить и в комплексе использовать оба подхода, а также учитывать происходящие в настоящее время политические и экономические события, оказывающие влияние как на внутренние, так и на внешние факторы деятельности каждого предприятия и соответственно обуславливающие рост или снижение цен на эти финансовые инструменты. Они способны обрабатывать огромные объемы данных, выявлять скрытые закономерности и создавать прогнозные модели, что открывает новые возможности для повышения точности анализа и принятия инвестиционных решений.

Искусственный интеллект (ИИ) дает целый ряд возможностей и преимуществ при решении сложных социально-экономических задач:

1. ИИ может быстро обработать огромные объемы данных и выявить скрытые закономерности, что может привести к более точным и значимым с практической точки зрения результатам.
2. Технологии ИИ позволяют использовать в принятии решений современные продвинутые инструменты, данные и информацию, которые дадут возможность

реализовывать данные процессы на новом качественном уровне.

3. ИИ и технологии больших данных могут применяться для обнаружения и предотвращения угроз безопасности, в сфере интернет-технологий и незаконных финансовых операций, осуществляемых в электронной среде.
4. Технологии ИИ позволяют сформировать механизмы сбора и обработки данных с фондового рынка, которые могут обеспечить оперативное и точное реагирование на изменения котировок финансовых инструментов и правильность принятия инвестиционных решений.
5. ИИ может быть использован для разработки оптимальных инвестиционных стратегий с учетом размеров вкладываемых капиталов и рисков, на которые готов пойти инвестор. Технологии ИИ позволяют сформировать алгоритмы автоматического выполнения принятых инвестиционных стратегий.

«Несмотря на перечисленные выше очевидные преимущества, точность прогнозов на основе нейросетей по фондовому рынку колеблется у отметки 50%, а всё потому, что фондовый рынок сложная органическая структура: факторы, которые значимы в один промежуток времени, теряют свою значимость в другой» [1].

В тоже время, согласно источнику [2] хедж-фонды создающие инвестиционные портфели на основе ИИ значительно превосходят совокупную доходность, показываемую традиционными методы инвестирования. Так исследование, проведенное аналитической компанией Cerulli Associates показало, что хедж-фонды Европейского союза, под управлением ИИ за период с 2016 по 2019 гг достигли совокупную доходность в 34% за три года, что почти в 3 раза превышало средний показатель по отрасли за тот же период.

Для формирования высокодоходных инвестиционных фондов хедж-фонды используют инструменты обработки естественного языка (NLP) посредством модели глубокого обучения на обработке актуальных социально-экономических и общественно-политических новостей для получения NLP-моделей прогнозирования стоимости различных финансовых инструментов с высоким уровнем точности. Эти модели в некоторых исследованиях демонстрировали средние ошибки прогнозирования менее 1-го % по целому ряду одних из наиболее ликвидных акций [3].

3. Результаты и обсуждение

К ограничениям в использовании технологий ИИ при работе на фондовых рынках и выстраивании финансовых стратегий, можно отнести:

- ошибки в прогнозах из-за некорректного ввода исходных данных;
- необходимость применения высокопроизводительной вычислительной техники для получения адекватных результатов;
- беспристрастность инструментов ИИ, что с одной стороны повышает качество аналитической работы, но с другой стороны может ухудшать правильность восприятия, в значительной степени, субъективного новостного фона, а также репутации компаний.

К методам ИИ, которые применяются в настоящее время для работы на фондовых рынках можно отнести [3]:

1. Анализ котировок акций, облигаций, фьючерсов, анализ финансовых данных, выявление тенденций, на основе интерпретации больших данных с помощью алгоритмов ИИ и машинного обучения.
2. Автоматические торговые алгоритмы «роботы» в режиме «реального времени» осуществляют мониторинг котировок акций, облигаций, валют и пр. финансовых инструментов и на этой основе осуществляют автоматические покупки и продажи данных инструментов в зависимости от полученных в ходе мониторинга торговых сигналов.
3. ARIMA — методика прогнозирования рыночных изменений на основе отслеживания и учета колебаний цен на финансовые инструменты на основе выявления исторических тенденций ценовых рядов и фактора сезонности. Наилучшим образом этот метод подходит для прогнозирования краткосрочных колебаний ценовых уровней.
4. Метод опорных векторов – модель ИИ, которая обучается на реальных данных, поступающих из торговой системы. При возможности обеспечить модели полноценный набор больших данных метод дает один из наиболее высоких уровней точности.
5. Метод «Случайного леса» — это алгоритм позволяющий использовать методы ИИ для анализа вероятности и тестирования моделей прогнозирования,

определенных событие, которые могут влиять на принятие инвестиционных решений.

Рассмотрим дополнительные методы, их преимущества и недостатки в таблице 1 ниже [4]:

Таблица 1. Сравнительный анализ нейросетевых моделей.

Метод машинного обучения	Преимущества	Недостатки
CNN: Сверточная нейронная сеть	+ Независимость влияния алгоритмов к повороту и сдвигу	- Продолжительное время обучения (несколько дней и более) для нейронной сети с числом слоёв свёртки свыше двух
	+ Обучается при помощи классического метода обратного распространения ошибки, при этом могут быть задействованы и другие методы обучения сети	- Пригодна, по большей части, только для распознавания изображений
	+ Использование ядер свёртки дает, в сравнении с полносвязанной нейросетью, меньшее количество настраиваемых весовых коэффициентов, что приводит к уменьшению времени и вычислительных ресурсов на обучение сети	- Большая вероятность переобучения сети при недостаточном количестве примеров при обучении с учителем
EMD: Метод эмпирической модовой декомпозиции	+ Позволяет проецировать нестационарный сигнал на частотно-временную плоскость с использованием монокомпонентных сигналов, что делает его адаптивным по своей природе	- Определение экстремумов изображения (или поверхности) и выбор метода интерполяции для применения на множестве точек рассеяния
	+ Не требует заранее заданного набора математических функций	- Частое появление смешивания мод
	+ Может обрабатывать адаптивные нестатические сигналы	
KNN: Метод k ближайших соседей	+ Алгоритм прост и легко реализуем	- Всегда нужно определять оптимальное значение k.
	+ Не чувствителен к выбросам	- Алгоритм работает значительно медленнее при увеличении объема выборки, предикторов или независимых переменных
	+ Нет необходимости строить модель, настраивать несколько параметров или делать дополнительные допущения	- Большие вычислительные затраты во время выполнения
LSTM: Долгая краткосрочная память	+ Способность анализировать и использовать взаимодействия и закономерности, существующие в данных, посредством процесса самообучения	- Отсутствует механизм индексации памяти при записи и чтении данных. Количество ячеек памяти привязано к размеру рекуррентных весовых матриц
	+ Делает хорошие прогнозы, потому что анализирует взаимодействия и скрытые закономерности в данных	
	+ Хорошо запоминает информацию в течение длительного времени	

Эта таблица помогает определить различные методы машинного обучения, их плюсы и минусы, а также выбрать наиболее подходящий метод для обучения ИИ в конкретных условиях и сетях. Также методы, приведенные в таблице, позволяет выявить тренды в развитии нейросетевых моделей, такие как увеличение глубины сетей, применение механизмов внимания и архитектуры трансформеров.

Что касается непосредственно процесса прогнозирования с помощью ИИ, он может включать в себя несколько этапов проведения подобной работы (рисунок 1).

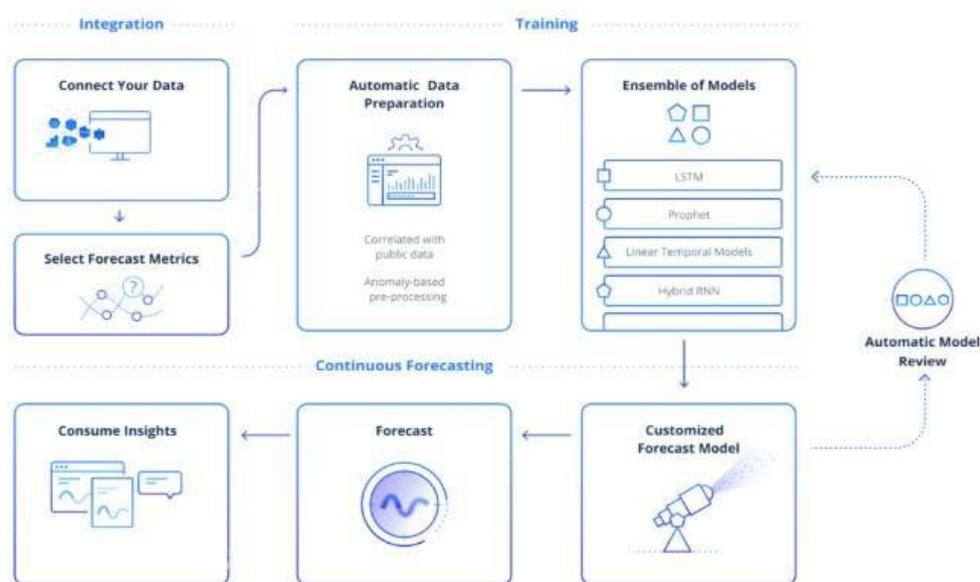


Рисунок 1. Процесс прогнозирования котировок финансовых инструментов с помощью ИИ.

Рассмотрим каждый этап подробнее:

1. Сбор и предварительная подготовка данных. В рамках данного этапа необходимо сформировать механизмы автоматической выгрузки данных о котировках акций, а также прочих данных и новостном фоне, которые могут быть необходимы для выставления моделей прогноза котировок.
2. Отбор наиболее приемлемых торговых индикаторов, определение значимости и типа новостей, необходимых для проведения прогноза. Оценка желаемых интервалов прогноза и необходимых для этого ретроспективных данных.
3. Разработка и включение в алгоритм механизмов по автоматизации предварительной подготовке данных, необходимых для осуществления

прогнозов. Важнейшей задачей на данном этапе является удаление факторов, которые не имеют отношения к прогнозу.

4. Обучение модели – проверка точности результатов, полученных разными методами прогноза и выбор метода, обеспечивающего наиболее полное соответствие фактическим значениям. Автоматизация полученных методик.
5. На основе оценки адекватности моделей, построенных на этапе 4, выбираются методы и модели, обеспечивающие прогнозирование котировок и финансовых результатов с наиболее высоким уровнем соответствия фактическим данным и текущей задаче прогнозирования.
6. Далее реализуется серия частных проверок точности, составленных моделей ИИ и больших данных для прогнозирования котировок финансовых инструментов, с которыми предполагается работа. При необходимости проверке могут осуществляться несколько раз.
7. Далее осуществляется непосредственно сам прогноз котировок отобранных для работы финансовых инструментов, формирование оптимальной инвестиционной стратегии и закупка первого пула финансовых инструментов в рамках реализации сформированной стратегии.
8. Автоматизация процесса отслеживания выполнения инвестиционной стратегии, ее постоянной корректировки и обновления в зависимости от изменения рыночной ситуации. Формирование специальных информационных панелей, графиков, отчетов и оповещений, необходимых для постоянного отслеживания эффективности работы торговой программы и оценке эффективности инвестиционной стратегии.

Вместо самостоятельной разработки программных кодов описанных выше алгоритмов, рациональнее всего использовать существующие и уже проверенные библиотеки ИИ. Существует множество готовых, поддерживаемых и простых в использовании библиотек, доступных для нескольких языков и платформ.

Так же, при выборе модели нужно определить цели прогноза: для чего используется модель? Хотим ли мы прогнозировать цену акции, объем торгов, изменение индекса или другие показатели? Из всех доступных данных необходимо выбрать показатели, которые, по мнению аналитика, имеют наибольшее влияние на целевой показатель. Например, для прогнозирования цены акции можно использовать показатели

прибыли, дохода, долга, индекса VIX (волатильность) и др. После выбора моделей и показателей прогноза необходимо собрать данные, опираясь на которые, искусственный интеллект выдает определенные данные. Чтобы облегчить этот процесс, можно использовать API для сбора данных из различных источников: финансовые площадки (Yahoo Finance, Google Finance), новостные агентства, социальные сети, индексы (S&P 500, NASDAQ), данные о макроэкономических показателях (ВВП, инфляция). Далее нужно откорректировать собранные данные: объединить данные из разных источников в единый датасет, устранить дубликаты, некорректные значения, пропущенные данные и др. Для бесперебойной работы нужно привести данные к единому формату (например, CSV). На основе точности каждой модели выбираются только самые эффективные из них для создания пользовательской модели для текущей задачи прогнозирования.

Далее после создания первоначальной настроенной модели нужно часто проводить проверки точности и при необходимости повторить предыдущие этапы обучения. Далее можно использовать пользовательскую модель и предоставленные данные для составления прогноза в режиме реального времени, а результаты сохраняются для будущего использования. Наконец, прогнозы отображаются в виде информационных панелей, отчетов или оповещений, так что ими может воспользоваться любой человек, независимо от его технических знаний.

Важно оценить результаты прогнозирования искусственным интеллектом для того, чтобы понять, верно ли проведены все вышеупомянутые этапы. Для этого можно применить следующие методы оценки (рисунок 2):

- средняя абсолютная процентная ошибка (MAPE):

$$MAPE = 100\% * \frac{1}{\text{Количество позиций}} * \sum \frac{|\text{Факт} - \text{Прогноз}|}{\text{Факт}}$$

- средняя абсолютная ошибка (MAE):

$$MAE = \frac{1}{\text{Количество позиций}} * \sum |\text{Факт} - \text{Прогноз}|$$

- средняя квадратичная ошибка (RMSE):

$$RMSE = \sqrt{\frac{1}{\text{Количество позиций}} * \sum (\text{Факт} - \text{Прогноз})^2}$$

Рисунок 2. Показатели оценки прогнозирования.

Средняя абсолютная ошибка (MAPE) – оценивается в процентах и используется для оценки величины ошибок в сравнении с фактическими значениями показателей.

Средняя абсолютная ошибка (MAE) представляет собой меру ошибок между парными наблюдениями одного и того же явления и применяется для прогнозирования спроса, анализа временных рядов и пр.

Средняя квадратичная ошибка (MSE) — это функция потерь, которая используется в линейной регрессии в качестве показателя эффективности.

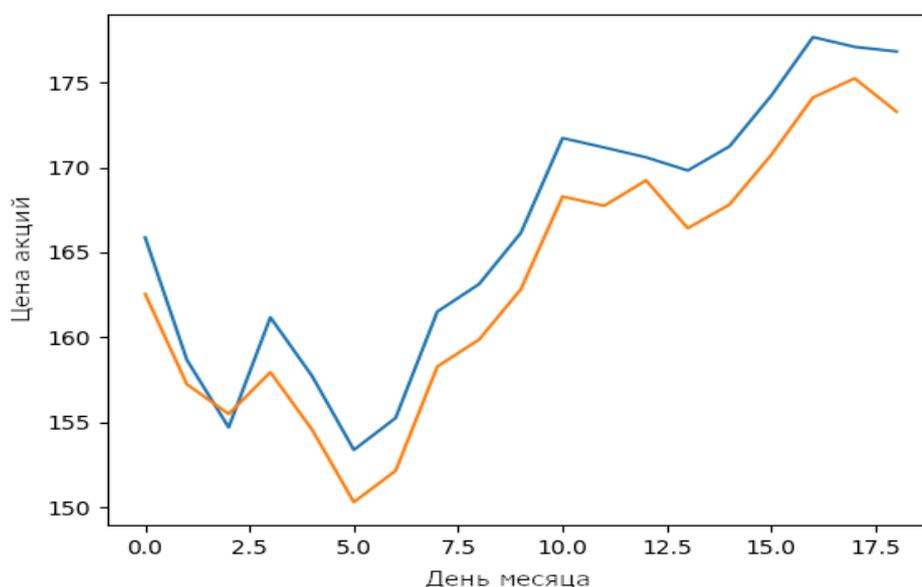


Рисунок 3. Прогноз модели LSTM на цены акций Apple в феврале.

Рассмотрим пример, когда прогноз цен акций поручили программному продукту с заложенным в него ИИ. Было протестировано 5 алгоритмов машинного обучения: линейная регрессия, градиентный бустинг, LSTM (Long Short-Term Memory) адаптивная авторегрессионная модель и логистическая регрессия. При этом использовались разные наборы данных и прогнозирование разных параметров. Наилучшие результаты были получены с помощью модели LSTM (рисунок 3) [5-7].

В качестве исходных данных были выбраны следующие показатели: поквартальные финансовые показатели компании: чистая и валовая прибыль, ликвидность, задолженность, EBITDA, рентабельность, активы, фундаментальные

мультипликаторы, коэффициенты Шарпа, Сортино, интегральный риск, волатильность и т. п.

Для простой LSTM модели без оптимизации — это очень хороший результат. Он показывает, что нейронные сети и модели машинного обучения способны строить сложные устойчивые связи между параметрами.

4. Заключение

Итак, искусственный интеллект и машинное обучение имеют огромный потенциал для улучшения анализа фондового рынка, но их применение требует тщательной проработки и оценки как преимуществ, так и ограничений. В будущем искусственный интеллект и машинное обучение, вероятно, станут неотъемлемой частью инвестиционного процесса, но инвесторам необходимо быть осторожными и критически оценивать результаты моделей ИИ. Важно продолжать исследования в области ИИ и машинного обучения для улучшения точности и эффективности моделей, разрабатывать этические и регуляторные рамки для использования ИИ в финансовой индустрии, обучать инвесторов основам использования данных методов в инвестировании.

Список литературы

1. Умное железо: как ИИ помогает инвесторам // БКС Экспресс: сайт. – URL: <https://bcs-express.ru/novosti-i-analitika/umnoe-zhelezo-kak-ii-pomogaet-investoram> (дата обращения: 01.06.2024).
2. Сергеев В.А. Использование нейросетей в прогнозировании фондового рынка / В.А. Сергеев // Azimuth of Scientific Research: Economics and Administration. – 2018. – Т. 7. – № 4(25). – URL: <https://cyberleninka.ru/article/n/ispolzovanie-neyrosetey-v-prognozirovanii-fondovogo-rynka/viewer> (дата обращения 06.06.2024).
3. Нешина К.С. Прогнозирование стоимости акций на основе нейросетевых моделей: дис. канд. экон. наук: 38.04.01. – Томск, 2022. – 150 с.
4. Машинное обучение: прогнозируем цены акций на фондовом рынке // Хабр: сайт. – URL: <https://habr.com/ru/companies/netologyru/articles/428227/> (дата обращения: 01.06.2024).
5. Агнон Х.О. Прогнозирование цены акций с использованием машинного обучения / Х.О. Агнон // Инновационная наука. – 2021. – № 6. – URL:

<https://cyberleninka.ru/article/n/prognozirovanie-tseny-aktsiy-s-ispolzovaniem-mashinnogo-obucheniya> (дата обращения: 01.06.2024).

6. Ways to Add Artificial Intelligence to an Existing Application // CodeProject: сайт. – URL: <https://www.codeproject.com/Articles/5330374/5-Ways-to-Add-Artificial-Intelligence-to-an-Existi> (дата обращения: 01.06.2024).
7. Random Forest Algorithm in Machine Learning // GeekforGeeks: сайт. – URL: <https://www.geeksforgeeks.org/random-forest-algorithm-in-machine-learning/> (дата обращения: 01.06.2024).