Санкт-Петербургский государственный университет

РАКШАЕВ Алдар Цыдендамбаевич

Выпускная квалификационная работа

Оптимальные стратегии на фондовом рынке в условиях нестабильной экономической ситуации

Уровень образования: магистратура

Направление *01.04.02 «Прикладная математика и информатика»*

Основная образовательная программа *ВМ.5505.2020 «Математическое и информационное обеспечение экономической деятельности»*

Научный руководитель:
Кандидат физико-математических наук,
старший преподаватель кафедры
математической теории игр
и статистических решений
Кумачева Сурия Шакировна
Рецензент:
Вильчевский Евгений Никитич

Санкт-Петербург

2022

Оглавление

Введение

Актуальность выпускной квалификационной работы. В последнее время произошло множество событий, способствовавших дестабилизации фондового рынка, таким образом, тематика исследования, которому посвящена данная выпускная квалификационная работа, актуальна потому, что необходим поиск эффективных стратегий и точных прогнозов в условиях нестабильности рынка от непредсказуемых внешних факторов.

В первую очередь, это пандемия коронавируса COVID-19, которая послужила причиной изменения спроса и предложения в большинстве областей. Многие производители товаров и услуг столкнулись со сложностями с производством и логистикой вследствие карантина и локдауна. С другой стороны, закрытие торговых точек стало причиной невозможности осуществления финансовых операций на местах.

Другой очевидной причиной нестабильности может служить сложная политическая ситуация. 24 Февраля 2022 года началась специальная военная операция Российской Федерации, что послужило причиной введения жестких санкций западных государств, приведших к падению котировок на биржах. В тот же день Индекс Московской Биржи упал на 45,49% - достигнув уровня 2015 года (1684,95 пункта).

Научная новизна работы. Модель Бокса-Дженкинса разрабатывалась в 70-х годах, когда экономический рынок был относительно стабилен. Сейчас, в условиях стремительно меняющегося рынка, представляется интересным оценить результаты ее применения с учетом корректировки параметров результатами анализа текущей ситуации фондового рынка.

Построение стратегии поведения игрока-инвестора в такой среде с помощью инструментария теории принятия решений и апробация данной стратегии на реальных данных с помощью приложения, созданного на базе библиотеки Python, составляет еще один интерес к данному научному исследованию. В условиях неопределенности должны существовать такие ориентиры, которыми бы мог руководствоваться субъект – в контексте данной задачи – лицо, принимающее решение (ЛПР).

Как уже известно, исторически экономические процессы имеют цикличный характер. И за развитием всегда будет следовать спад экономики. Чтобы выдержать такие переменчивые внешние обстоятельства, необходимо разработать методику поиска оптимальной стратегии поведения как наилучший ответ на условия рынка.

В 2020 году мировая экономика сократилась на 4,3 процента, что более, чем в два с половиной раза, больше, чем во время мирового финансового кризиса 2009 года. Восстановление после пандемии будет зависеть не только от масштабов стимулирующих мер и быстрого внедрения вакцин, но и от качества и эффективности этих мер в целях повышения устойчивости к будущим потрясениям [1].

С начала пандемии и карантинного режима число инвесторов в России значительно повысилось, особенно частных инвесторов. Параллельно происходит изменение структуры финансового рынка: все большее влияние на население оказывают электронные и интернет-корпорации. В связи с пандемией, многие компании, работающие в сфере общественного питания, развлечений и туризма, продолжают терпеть убытки.

В такой нестабильной ситуации на фондовой бирже составление прогнозов призвано помочь специалистам при разработке оптимальных алгоритмов дальнейшего развития рынка каждого конкретного государства в частности и мировой экономики в целом.

Популярным и широко используемым статистическим методом для прогнозирования временных рядов является модель ARIMA. ARIMA - это аббревиатура от Auto Regressive Integrated Moving Average. Это класс модели, который фиксирует набор различных стандартных временных структур и данных временных рядов. Этот подход к моделированию позволяет выбирать тенденции, используя авторегрессию, скользящие средние и сезонные различия в расчетах [2].

Текущее исследование посвящено решению задачи прогнозирования динамики фондового рынка, основанному на методике применения этой модели, а также поиску оптимальной стратегии игрока-инвестора в среде, для которой получен прогноз.

Ключевые слова. фондовый рынок, индекс Московской биржи, индекс РТС, индекс голубых фишек, временные ряды, нестационарность, модель Бокса-Дженкинса, эконометрика, теория принятия решений в условиях неопределенности и риска, теория игр, оптимальные стратегии.

Постановка задачи

В условиях экономической нестабильности для описания и анализа любых динамических процессов, связанных с фондовым рынком, важнейшее значение имеет оценка рисков, сопряженных с этим процессом. То есть, задачу поиска оптимальной стратегии можно интерпретировать следующим образом: нужно найти такую стратегию, чтобы риски при ее применении были минимальными. Поиск оптимальных стратегий на фондовом рынке был и остается основной целью множества аналитиков-экономистов.

В настоящее время создано много программных средств, позволяющих производить расчеты и давать некоторые рекомендации при продаже или покупке товаров на фондовом рынке. К сожалению, большинство доступных программных продуктов не дает нужный результат, позволяющий аналитику с большой вероятностью оценить происходящую на фондовом рынке ситуацию и принять эффективное решение.

Для проведения анализа сложных структур с помощью методов прикладной статистики и эконометрики одним из самых популярных программных средств является Python – этот интерпретируемый объектно-ориентированный язык имеет открытый исходный код и динамическую семантику. Предпочтение ему отдается в силу его интерпретируемости, которая сильно упрощает тестирование и перемещение с платформы на платформу небольших блоков кода. [3]

Глобальная цель данного исследования: построение алгоритма поиска оптимальной стратегии. Математически это задача о принятии решений в условиях неопределенности, сопровождаемая решением задачи прогнозирования.

Предметная цель: выявление зависимостей поведения субъектов финансового рынка от результатов прогнозных значений.

Исходя из постановки задачи принятия решений, рассмотрим в качестве ЛПР игрока, стремящегося найти оптимальную стратегию на фондовом рынке. В данных условиях можно выделить следующие задачи:

1. Изучение современных методов статистического прогнозирования;
2. Применение модели Бокса-Дженкинса для прогнозирования данных временного ряда;
3. Описание процесса нахождения оптимальной стратегии;
4. Апробация модели на реальных данных.

Объектом данного исследования является фондовый рынок.

Предметом исследования выступают котировки Индекса Московской Биржи, РТС и Голубых фишек.

Обзор литературы

Существует огромное множество методов прогнозирования, и каждый применяется для разных ситуаций с конкретной системой условий и ограничений. Прогнозирование и анализ полученных результатов являются основными темами в области разработки оптимальных стратегий. В работе [2] рассматриваются методы анализа коротких временных рядов. Подробно обсуждаются проблемы, связанные с анализом временных рядов в условиях малого числа наблюдений, возможные подходы к прогнозированию коротких временных рядов и методы прогнозирования с привлечением экспертной информации. Материал содержит изложение основных достижений в данной области.

Теоретические и прикладные аспекты эконометрики систематически изложены в [4]: тщательно разбираются и анализируются различные линейные регрессионные модели.

Методологическую основу одной из частей проводимого исследования составил метод Бокса-Дженкинса, введенный в использование в [5]. Примечательно, что в данном издании основное внимание уделено практическому применению моделей авторегрессии, в особенности программированию на ЭВМ, несмотря на то, что исторически впервые полное описание модели появилось именно в этом труде.

Немаловажный интерес представляет [6], в которой представлены статистические методы микроэкономики, макроэкономики и финансового анализа. Эконометрические модели всегда относятся к решению задач прогнозирования. Также разобраны различные примеры применения их на практике.

Условные обозначения

Цены акций, облигаций, производных финансовых инструментов, изменяясь во времени, образуют ряд. Любые финансовые данные, такие, как наборы последовательных значений показателя y_1, y_2, …, y_t, …, y_T, зафиксированных в равноотстоящие друг от друга моменты времени t=1, 2, …, T, так, что интервал (t, t+1) является постоянным, являются временным рядом. Такой ряд представляет собой дискретный временной процесс.

Случайный процесс, протекающий в некоторой системе S, называется процессом с дискретным временем, если переходы системы из одного состояния в другое происходит в заранее известные моменты времени t_0, t_1, …, t_k, которые называют шагами или этапами процесса. В промежутки времени между смежными шагами состояние системы не изменяется.

Математически это выражается следующей функциональной зависимостью:

$$y_t=f\,(y_{t-1},\, y_{t-2},\, \dots) + \varepsilon_t \tag{0.1}$$

где f – некоторая функция; ε_t – ошибка модели, представляющая собой «белый шум» - стохастический стационарный процесс, обладающий следующими свойствами:

- Математическое ожидание: $\mu_\varepsilon = 0$
- Дисперсия постоянна и равна: σ_ε^2
- «Белый шум» распределен нормально: $\varepsilon_t \sim N(0; \sigma_\varepsilon^2), t = 1, \dots, T$
- Элементы белого шума взаимно не коррелируемы cov (ε_i; ε_j) = 0, $i \neq j$
- Набор случайных переменных y_t, $t \in R$ - стохастический процесс.

Глава 1. Модель Бокса-Дженкинса

1.1. Экономические методы прогнозирования

Дадим определение фондовому рынку, обсудим, что собой представляет Московская Биржа, что такое индексы и котировки, которыми оперируют на данном рынке.

Итак, **фондовый рынок**, согласно определению РБК Инвестиции [7] - это место, где происходит торговля акциями, облигациями, валютами и прочими активами. Понятие рынка затрагивает не только функцию передачи ценных бумаг, но и другие операции с ними, такие, как выпуск и налогообложение. Кроме того, он позволяет устанавливать справедливое ценообразование.

Рынок ценных бумаг имеет определенные признаки:

- У него всегда есть фиксированная торговая площадка, например, фондовый рынок Московской биржи;
- Обязательно наличие специализированного механизма отбора товаров (активов), отвечающих определенным требованиям;
- Установлены торговые процедуры по времени и стандартам;
- Все оформление сделок централизованно;
- Деятельность всех участников рынка контролируется уполномоченными органами;
- Существуют официальные котировки активов.

Схематично это выглядит так, как представлено на рис.:

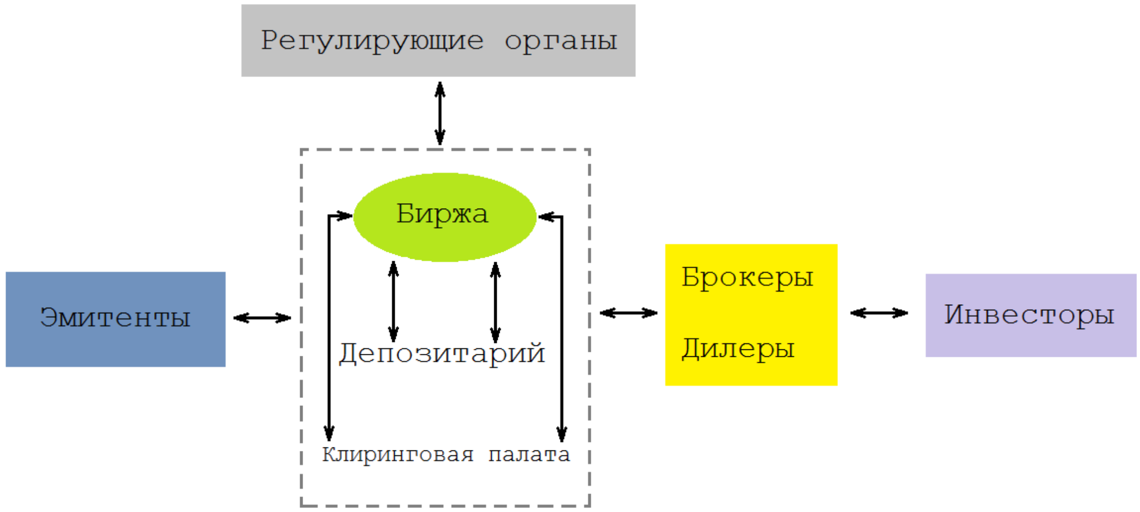


Рис. 1. Структура фондового рынка

Фондовый рынок является одним из самых популярных инвестиций, благодаря высокой ожидаемой прибыли. Однако, чем выше ожидаемая прибыль, тем выше подразумеваемый риск. Фондовый рынок, который был исследован ранее в различных работах [8], является достаточно компилированной средой, или доступной для математического анализа.

Согласно [9], существует три степени эффективности рынка, отражение каждой из которых сформулировано в следующих гипотезах. Одна из гипотез эффективного рынка говорит о том, что вся информация, которую можно получить, сразу же учитывается в рыночной цене ценной бумаги. По другой гипотезе эффективных рынков считается, что вся публичная информация сразу же отражается в цене, как она стала известна, но владельцы частной информации могут использовать эту информацию для получения прибыли. Третья гипотеза заключается только в том, что любая информация, полученная в результате изучения прошлой торговой истории, отражается в цене.

Существует два вида теоретических подходов для определения входной переменной для прогнозирования индекса фондового рынка с помощью нейронных сетей [10]. Первый вводит взаимосвязь между индексной ценой фондового рынка и другими макроэкономическими показателями. Второй вводит нелинейность в отношении между ценами на акции, дивидендами и объемом торгов.

Прогнозирование тенденций финансового рынка является одной из важнейших задач для инвесторов. С 70-х годов участникам финансового рынка рекомендовалось использовать стратегию «покупай и держи», которая основывалась на теории случайных блужданий. Считалось, что ряд содержит огромное количество скрытых закономерностей, которые пытались выявить с 30-х годов последователи технического анализа, созданного Р. Эллиотом. Хотя ряды выглядели случайными, но вполне допускали краткосрочное прогнозирование, что давало получение реального дохода. Но тот технический анализ при своей наглядности, универсальности и удивительной простоте использования, имел ряд недостатков. Массовое применение анализа размыло основные уровни, рынок часто отходил от специальных моделей, что урезало эффективность применения.

Те, кто обладает усовершенствованными методами выявления закономерностей в хаотических рядах, получают наибольшую прибыль, сделав максимально точный вывод относительно сложившейся ситуации.

Данные финансовых временных рядов являются более сложными, чем другие статистические данные, из-за долгосрочных тенденций, циклических колебаний, сезонных колебаний и нерегулярных движений. Прогнозирование таких сильно колеблющихся и нерегулярных данных обычно подвержено большим ошибкам. Поэтому разработка более реалистичных моделей для прогнозирования данных финансовых временных рядов для более эффективного и точного извлечения из них значимых статистических данных представляет большой интерес для исследований в области анализа финансовых данных.

В течение многих лет исследователями была проведена большая работа по разработке эффективных моделей для повышения точности прогнозирования [11]. Точность прогноза измеряется в результате расчета ошибки прогнозирования временного ряда. В результате в литературе появились различные модели прогнозирования временных рядов, такие как регрессионные, авторегрессионные модели, модели на основе генетического алгоритма, модели на опорных векторах, модели на нейронных сетях и т.д.

Временной ряд (динамический ряд или ряд динамики) – это последовательность, в которой наблюдаются некоторые признаки (случайные величины). Отдельное наблюдение называется уровнем ряда, обозначаемым через x_t (t = 1,2, ..., n), где n - число уровней. [12]

В общем виде при исследовании экономического временного ряда x_t выделяются несколько составляющих (аддитивная модель):

$$x_t = T + S + C + E, \text{где } (t = 1, 2, \dots, n) \qquad (1.1)$$

или мультипликативная модель:

$$x_t = T * S * C * E \qquad (1.2)$$

где T - это тренд, который определяет чистый эффект долгосрочного фактора, медленно меняющегося компонента, то есть долгосрочную тенденцию изменения свойств (например, рост населения, экономическое развитие или изменение моделей потребления и т.д.);

S – сезонная составляющая, отражающая экономические процессы в течение очень длительного времени (годы, иногда месяцы, недели и т. д. Например, при продаже товаров, перевозке пассажиров, в разное время года);

C является циклическим компонентом, отражающим повторение затянувшегося экономического процесса (например, волна экономической активности Кондратьева, демографическая «яма», и т.д.);

E является случайным компонентом и отражает влияние случайного фактора.

В отличие от E, первые три компонента T, S и C являются регулярными и нестандартными. Важнейшей классической задачей изучения временного ряда является выявление отклонения от основных тенденций развития процесса.

Основные этапы анализа временных рядов следующие:

1. Графическое представление и интерпретация поведения временных рядов.

2. Выбор или удаление регулярных (не случайных) компонент временного ряда (трендовый, сезонный и циклический компоненты).

3. Сглаживание и фильтрация (удаление низкочастотных и высокочастотных составляющих).

4. Исследование случайных компонентов временных рядов, а также составление и проверка математических моделей, соответствующих их определениям.

5. Построение прогноза развития процесса на основе временных рядов.

6. Изучение отношения между различными временными рядами.

Колебания фондового рынка являются результатом сложных явлений, эффект которых выражается в сочетании прибылей и убытков, которые появляются на графике временных рядов цены акций. Говоря о временных рядах, описывающих именно фондовые рынки, наиболее заметные изменения в структуре ряда – это тренд, периодические изменения и ежедневные изменения. Тренд – это идентифицируемая долгосрочная вариация во временных рядах фондового рынка, которая обычно прогнозируется экстраполяцией [13]. Периодические колебания следуют либо за сезонными моделями, либо за деловым циклом в экономике. Краткосрочные и ежедневные колебания появляются случайным образом и их трудно предсказать, но они часто являются источником прибыли и убытков при торговле акциями, особенно в случае дневных трейдеров.

Среди наиболее распространенных методов анализа временных рядов можно выделить: корреляционный и спектральный анализ, авторегрессионные модели и скользящие средние. В то же время необходимо учитывать принципиальное отличие наблюдаемой последовательности серий случайных выборок x (t = 1, 2, ..., n), $x_1, x_2, ..., x_t, ..., x_n$.

Во-первых, в отличие от элементов случайной выборки, члены временных рядов, как правило, статистически независимы.

Во-вторых, члены временного ряда распределяются неравномерно.

Один из сложных вопросов в прогнозировании временных рядов - как найти лучший алгоритм. В последние годы была разработана система рекомендаций для анализа временных рядов с использованием подхода метаобучения. Эта система выбирает лучший метод прогнозирования с учетом характеристик временных рядов. [14]

Чтобы выбрать модель для прогнозирования временного ряда, используются два метода: совокупный отбор и индивидуальный отбор. При совокупном отборе выбирается одна модель для прогнозирования всех временных рядов, а при индивидуальном отборе для каждого временного ряда выбирается модель с наилучшими показателями [3]. Чтобы выбрать лучший метод прогнозирования для каждого временного ряда, были разработаны два основных подхода: информационные критерии и эмпирическая точность. В первом подходе используются информационные критерии Акаике (AIC) и его расширения для выбора лучшей модели для определенного временного ряда. Во втором подходе, с помощью кросс-валидации и оценки ошибок, в качестве подходящей модели для временного ряда выбирается модель с наилучшими показателями [15].

Одним из наиболее существенных недостатков методологий выбора модели являются затраты на осуществление вычислений, поскольку обычно тестируются все данные. Если набор данных превышает определенный порог, вычислительная нагрузка является высокой. Эти проблемы побудили исследователей искать альтернативы построения прогнозных моделей.

Одной из альтернатив предыдущим подходам выбора модели прогнозирования временного ряда является создание системы рекомендаций, использующей алгоритмы машинного обучения, которые способны предложить наиболее мощный метод прогнозирования для конкретного временного ряда.

В качестве метода прогнозирования принято понимать комбинацию методов мышления, основанных на анализе прошлых тенденций (ретроспективный анализ) и способности искать шаблоны между переменными модели. В научном сообществе нет определенной классификации, и разнообразие методов прогнозирования временных рядов формально, но можно предложить следующую классификацию, которая более подробно описывает методы математического прогнозирования [16]:

1. Линейные и нелинейные методы.

Линейная модель является основой для построения прогноза, она определяет связь между переменными в форме линейной функции и имеет вид (1.3):

$$y = \beta_0 + \sum \beta_i X_i + \varepsilon_i \tag{1.3}$$

Наиболее распространенным примером линейных моделей является линейная регрессия, которая наиболее изучена в эконометрике. Такой дизайн не эффективен, но это не так. Линейные модели могут решить не только проблему предсказуемости, но и агломерацию и классификацию, которая активно используется страховыми компаниями.

Нелинейная модель определяет нелинейные отношения между зависимыми переменными и имеет вид (1.4) [17]:

$$y = f(X, \beta) + \varepsilon \tag{1.4}$$

2. Сосредоточенные (зависящие от одной переменной) или распределенные (зависящие от нескольких переменных) [15].

Сосредоточенные модели определяют динамику системы, состоящей из дискретных компонентов - системы линейных и нелинейных уравнений.

3. Детерминированные (предопределенные, уникальные) или стохастические (случайные).

Детерминированные модели – это такие алгоритмы, которые дают уникальные, заранее определенные результаты для любого источника. Стохастические модели – это уравнения для случайного процесса, которое широко распространено не только в экономике, но и в физике, биологии и медицине.

4. Статические или динамические.

Ключевым разрывом между статическими или динамическими моделями является момент времени построения связей. Для статических моделей все зависимости являются мгновенными и динамически различными.

5. Дискретные (конечные целочисленные модели) или непрерывные (трудоемкие процессы).

Большинство математических методов можно использовать только в стационарных случаях, редким исключением являются финансовые временные ряды, что значительно усложняет моделирование. Стационарные временные ряды со временем не меняют своих свойств и характеризуются следующими свойствами: [3]

1. Математическое ожидание постоянно (1.5):

$$\forall t E(y_t) = \bar{y} = const \tag{1.5}$$

Это означает, что существует предложение, что стационарный временной ряд не имеет тенденцию к росту или падению, а в среднем остается на одном и том же уровне.

2. Дисперсия временного ряда постоянна в любом разрезе (1.6).

$$\forall t D(y) = E(y_t - \bar{y})^2 = G^2(y) = const \tag{1.6}$$

Это означает, что колебания ряда во времени (что год назад, что сегодня) в среднем одинаковые.

3. Коэффициенты автокорреляции (p_k) с любым лагом k, являются постоянными (1.7): [3]

$$\forall k p_k = \frac{E(y_t-\bar{y})(y_{t+k}-\bar{y})}{\sqrt{E(y_t-\bar{y})}E(y_{t+k}-\bar{y})} = f_k \tag{1.7}$$

Это означает, что линейная зависимость индикатора от предыдущего одинакова.

Временной ряд, который не удовлетворяет по меньшей мере одной из вышеупомянутых характеристик, часто называют нефиксированным временным рядом. Проблема в том, что математические методы прогнозирования, как правило, обобщены, но они показывают плохие результаты по сравнению с компонентными переменными. Многие модели основаны на разложении сезонных, трендовых и случайных компонентов, но их определения имеют значительные отклонения из-за неоднородности ряда.

Прежде всего, при проектировании необходимо сделать серию стационарной. Есть несколько основных способов избежать нестабильности:

1. Интегрирование временных рядов.

Преобразование в один временной ряд включает в себя некоторые последовательные различия. Поэтому ряд называется суммой порядков k (1.8), когда разность порядка k постоянна.

$$X_t \approx I(k), \Delta^k x_t \text{ - стационарные} \tag{1.8}$$

Как показывает практика, не каждый финансовый временной ряд может быть преобразован в стационарный вид. Это связано с тем, что данные могут быть неполными или вероятными. Но метод часто используется для моделирования.

2. Итеративные процедуры.

Этот метод включает в себя «подсчет» возможных значений коэффициентов модели и выбор оптимальных значений, определяемых подстановкой. В этом методе параметры определяют заданную надежность, которая может быть скорректирована в соответствии с требованиями задачи моделирования.

Временные ряды имеют измерения, которые затрудняют прогнозирование. В математическом моделировании необходимо нормализовать переменные. Нормализация – это форма обработки временных рядов, в результате которой ряд становится безразмерным. Различают следующие методы нормализации:

1. Линейная нормализация [19].

1.1. Нормализация в пределах [0,1]:

$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{1.9}$$

1.2. Нормализация в пределах [-1,1]:

$$x_i = \frac{x_i - \bar{x}}{x_{max} - x_{min}} \tag{1.10}$$

1.3. Нормализация в пределах [-1,0]:

$$x_i = \frac{x_i - x_{max}}{x_{max} - x_{min}} \tag{1.11}$$

2. Нелинейная нормализация [16].

Экспоненциальные или логарифмические функции используются для нормализации нелинейных данных, и их выбор зависит от задачи. Линейная нормализация является наиболее распространенным методом.

Нелинейные вычисления оценивают взаимосвязь между зависимой и независимой переменными. Есть два основных вида нелинейных моделей.

1. Линейная структура регрессионных моделей.

Наиболее распространенной моделью является полиномиальная регрессия, основанная на ретроспективных данных (например, 1.12).

$$y = a + b_2 x + b_2 x^2 \tag{1.12}$$

Полиномиальная регрессия используется для моделирования ряда временных тенденций и в целях прогнозирования.

2. Существенно нелинейные регрессионные модели.

Такие модели называются нелинейными регрессионными моделями и не могут быть сведены к линейным (например, 1.13).

$$y = e^{-b_i * x} + \varepsilon \tag{1.13}$$

Здесь основная задача - оценить коэффициенты нелинейной модели. Такой процесс называется нелинейным расчетом и включает следующие методы:

1. Метод наименьших квадратов [3]. Метод основан на минимизации суммы квадратов отклонений функции от истинного значения. Суть метода можно выразить формулой (1.14):

$$\sum_i e_i^2 = \sum_i (y_i - f_i(x))^2 \rightarrow min \tag{1.14}$$

2. Построение функции потерь [3]. Такая функция определяет меру разницы между истинным значением параметра и его вычислением.

3. Метод взвешенных наименьших квадратов [3]. Этот метод хорошо работает, когда остаточное распределение является однородным.

4. Метод максимального правдоподобия [3]. Расчет метода основан на вычислении неизвестного коэффициента путем увеличения функции вероятности. Все связанные данные выборки считаются сохраненными в функции вероятности.

5. Алгоритмы минимизации функций [3]. Алгоритмы основан на вычислении второй производной функции потерь, которая используется для поиска наименьшего размера.

6. Симплекс метод. [3] Алгоритм предназначен для подсчета решения задачи линейного программирования и позволяет найти оптимальное решение за несколько шагов и определить, отсутствует оно или нет.

1.2. Модели авторегрессии – проинтегрированного скользящего среднего

В целом модели для данных временных рядов могут иметь множество форм и представлять собой различные модели стохастических процессов. В литературе [4] широко используются две модели линейных временных рядов, а именно: модели с авторегрессии (АР) и скользящего среднего (КС).

Объединяя эти две модели, получим авторегрессию со скользящим средним (ARMA) и авторегрессию с интегрированным скользящим средним (ARIMA), которые ранее были изучены в [2], [13], [6]. Модель авторегрессионная фракционно-интегрированная скользящего среднего (ARFIMA) обобщает модели ARMA и ARIMA [2]. Для прогнозирования сезонных временных рядов используется вариация модели ARIMA, то есть Сезонная авто регрессия с интегрированным скользящим средним (-SARIMA) [13].

Модель ARIMA и ее различные варианты основаны на знаменитом принципе Бокса-Дженкинса [6], поэтому они также широко известны как модели Бокса-Дженкинса. Линейные модели привлекли большое внимание, благодаря своей относительной простоте в понимании и реализации.

Нелинейные модели подходят для прогнозирования изменений волатильности в экономических и финансовых временных рядах. Учитывая эти факты, в литературе предлагаются различные нелинейные модели: условной гетероскедастичности (ARCH) и ее разновидности, такие как: обобщенная ARCH (GARCH), экспоненциально обобщенная ARCH (EGARCH), модель нелинейной авторегрессии (NAR), модель нелинейного скользящего среднего (NMA), авторегрессия со скользящим средним [18].

Наиболее распространенными нелинейными моделями являются ARMA (2.1) или ARSS (в среднем с движениями авто реверсии), которые объединяют две простые модели: AR (модель авторегрессии) и MA (модель скользящего среднего): [3]

$$X_t = c + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \sum_{i=1}^{p} \varphi_i X_{t-i}, \varepsilon_t = N(0, \sigma^2) \qquad (2.1)$$

Здесь коэффициент p - это порядок сегмента регрессии, а коэффициент q - это скользящее среднее.

В модели AR(p) предполагается, что новое значение переменной будет линейной комбинацией p прошлых наблюдений и случайной погрешности в сочетании с постоянными сроками. Математически модель AR(p) может быть выражена как: [19]

$$y_t = c + \sum \varphi_i y_t - p_i \qquad (2.2)$$

$$i = 1 + Et = c + \varphi 1 yt - 1 + \varphi 2 yt - 2 + ... + \varphi pyt - p + \varepsilon t \qquad (2.3)$$

Период $t, i\ (i = 1,2, \dots, p)$ - параметры модели и c - константа. Целочисленная константа p называется порядком прохождения модели. Иногда константа опускается для простоты.

Подобно тому, как AR(p) модель регрессирует по отношению к прошлым значений серии, MA(q) модель использует прошлые значения ошибок, как объясняющих переменных. Модель MA(q):

$$yt = \mu + \sum \theta j \varepsilon t - iq \qquad (2.4)$$

$$j = 1 + \varepsilon t = \mu + \theta 1 \varepsilon t - 1 + \theta 2 \varepsilon t - 2 + ... + \theta p \varepsilon t - q + \varepsilon t \qquad (2.5)$$

где μ - среднее значение ряда; $\theta j(j1,2,...q)$- параметры модели; q - параметры модели порядок прохождения модели.

Таким образом, концептуально модель скользящего среднего представляет собой линейную регрессию текущего наблюдения временного ряда против случайных толчков одного или более предыдущих временных рядов. Подгонка MA-модели к временным рядам сложнее, чем подгонка AR, так как в первом случае случайные члены ошибки не будет.

Модели с авторегрессией (AR) и скользящим средним (MA) могут быть эффективно объединены вместе. Для формирования общего и полезного класса моделей временных рядов, известных как модели ARMA.

Математически модель ARMA (p, q) представлена как:

$$yt = c + \varepsilon t + \sum \phi iyt - ip \tag{2.6}$$

$$i = 1 + \sum \theta j \varepsilon t - jq \tag{2.7}$$

$$\text{j=1} \tag{2.8}$$

Здесь модель упорядочивает p, q согласно p авторегрессии и q условиями скользящего среднего.

Обычно модели ARMA изменяются с помощью нотации лагового оператора. Оператор запаздывания или обратного смещения определяется как $L_{yt} = y_t - 1$

Полиномы лагового оператора или лаговые полиномы используются для представления моделей ARMA следующим образом: AR (p) модель:

$$\varepsilon t = \phi(L) yt \tag{2.9}$$

MA модель:

$$yt = \theta(L) \varepsilon t \tag{2.10}$$

ARMA модель:

$$\phi(L) yt = \theta(L) \varepsilon t \tag{2.11}$$

Важным свойством процесса AR (p) является оборачиваемость, то есть процесс AR (p) всегда может быть записан в виде процесса MA (∞).

Если MA (q) процесс должен быть инвертированным, то все корни уравнения θ = 0 должны лежать вне единичного круга. Это условие известно, как условие инвертованости для процесса MA. [19]

Обычно стохастический процесс, управляющий временными рядами, неизвестный, и, следовательно, невозможно определить фактические или теоретические значения АКФ и ЧАКФ: автокорреляционную функцию и частную автокорреляционную функцию.

Методология построения модели класса ARMA заключается в следующем:

1. Постулирование.

На данный момент выбирается категория модели.

2. Идентификация.

Следующим шагом является определение подкатегорий моделей. Опция подкатегории заключается в выборе конкретного коэффициента модели.

3. Оценка.

Значения параметров модели рассчитываются методом наибольшей вероятности.

4. Диагностика.

Последний шаг - проверить соответствие шаблона, изучить его свойства и принять соответствие.

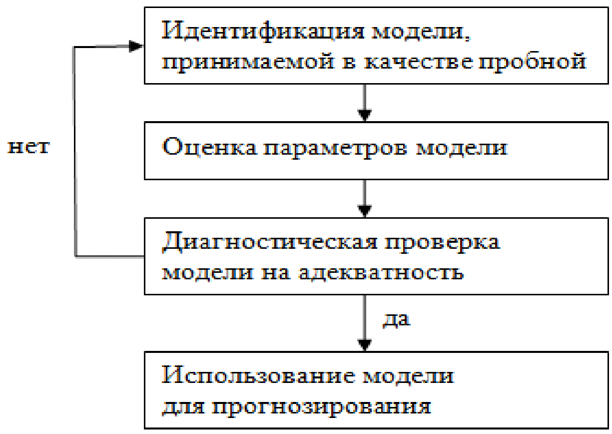


Рис. 2. Общая схема выбора модели

Модель ARMA, описанная выше, может использоваться только для данных стационарных временных рядов. Однако, на практике многие временные ряды относятся к социально-экономическим и проявляют нестационарное поведение. Временные ряды, содержащие тренды и сезонные модели, являются также нестационарными по своему характеру.

Чтобы решить проблему временных рядов, можно добавить интеграцию в модель для ее решения. Результатом является ARIMA (2.12) (интегрированная модель скользящего среднего). Эта модель представляет собой более широкую версию ARMA, которая может быть преобразована в нестационарную с учетом различий во временных рядах. В модели коэффициент "d" указан в понимании порядка разности.

В моделях ARIMA нестационарные временные ряды становятся стационарными путем нанесения конечного значения различия в точках данных.

Алгоритм выбора модели ARIMA (p, d, q):

1. Вычисляем значение автоковариации в зависимости от лага для данного ряда (d=0) и разностных рядов первого и второго порядка (d=1,2) и выбираем такой порядок d, при котором величина автоковариации резко уменьшается при небольшом лаге.
2. С помощью критериев Акаике (AIC) и байесовского критерия Шварца (SBC) выбираем наиболее подходящие значения p и q.
3. Оцениваем параметры моделей для данного ряда при выбранных значениях p, d, q методов наименьших квадратов.
4. Производим прогноз на выбранном горизонте прогнозирования с использованием выбранных параметров модели. [18]

Полезным обобщением моделей ARIMA является авторегрессионная модель с мелко интегрированным скользящим средним, которая допускает не числовые значения разницы коэффициента «d», то есть может быть получен дробно (2.17). Отрицательное значение объясняется расширением оператора в степенной ряд. Эта модель называется ARFIMA (модель с скользящим средним делением). [19]

$$\Delta^d = (1-L)^d = \sum_{k=0}^{\infty} \Pi(d-j)\frac{(-1)^k}{k!}L^k \tag{2.17}$$

ARFIMA имеет полезное применение в моделировании временных рядов с большой памятью. В данной модели расширение срока должно быть сделано с использованием общей биномиальной теоремы.

В том случае, когда временные интервалы малы, модель биномиального ценообразования опционов [20] демонстрирует неплохое приближение. Эта модель строит дерево (или решетку) решений, потому что каждая ветвь делится на большее количество ветвей. Модель биномиального ценообразования отображает эволюцию рыночной стоимости опциона в шкале дискретного времени. Каждый узел в решетке представляет собой возможную цену опциона в данный момент времени. Оценка выполняется итеративно, начиная с каждого из конечных узлов. Она широко используется для нестационарных данных, таких как экономические ряды и временные ряды для котировки акций. [21]

Финансовые временные ряды являются сезонными, поэтому они не учитываются в вышеупомянутых схемах. Сезонность - это периодическое изменение временного ряда. Сезонные модели называются SARIMA [22] или ARIMA (p, d, q) (P, D, Q) s. Коэффициент «P» представляет собой последовательность SAR квартального компонента (P), коэффициент «D» представляет собой последовательность консолидации квартального компонента, коэффициент «Q» представляет собой последовательность квартального состава, SM - размер квартала (неделя, месяц и т. д.).

В этой модели сезонная дифференциация соответствующего порядка используется для удаления не стационарности временного ряда. Сезонная разница первого порядка – это разность между наблюдением данного времени и соответствующим наблюдением в предыдущем году и рассчитывается как

$$z_t = y_1 - y_t - s \tag{2.18}$$

Для месячных временных рядов s = 12 и для квартальных временных рядов s = 4.

Существует много способов определения сезонности, но наиболее распространенным является автокорреляционная функция. [16]

Автокорреляция (статистически) - это функция корреляции в разные моменты времени (2.19), а функция автокорреляции - это последовательность коэффициентов корреляции.

$$r = \frac{E[(y_t - \bar{y})(y_{t+k} - \bar{y})]}{\sqrt{E(y_t - \bar{y})E(y_{t+k} - y)}} \qquad (2.19)$$

Основным преимуществом этого метода является то, что результаты легко реализуются и интерпретируются.

1.3. Применение теории принятия решений на фондовом рынке

Принятие решений на фондовом рынке представляет собой «игру с природой» - теория принятия решений с оппонентом, не заинтересованным в каком-либо исходе. В качестве «природы» выступают значения прогнозов индексов котировок московской биржи, РТС и голубых фишек.

При составлении игры с природой первым делом следует начать с построения матрицы выигрышей. Причем, у второго игрока, называемого «природой» нет цели выиграть или проиграть, и, таким образом, задача может быть сформулирована как антагонистическая игра без противника.

Выработка оптимальной стратегии поведения может осуществляться с использованием методов уже упомянутой теории игр, причем, в соответствии с делением условий на две разновидности: моделирование производится, как по классическим принципам состязания двух разумных игроков (конкурентная борьба), так и по принципам, когда противника нет, как такового, т.е. игроку противостоят силы, не заинтересованные в выигрыше или проигрыше какой-то из сторон, есть только определенная среда или природа, со своим набором условий из множества условий. [23]

В играх с природой наиболее часто используются критерии Вальда, Сэвиджа и Гурвица [24]. Критерий Вальда - максиминный критерий и является пессимистическим [23]. Применяется в том случае, когда игрок сомневается в надежности статистических вероятностей состояний природы и хочет застраховать себя от неожиданных проигрышей. В нашем случае критерий Вальда выглядит следующим образом:

$$V_W = \max_i \min_k g_{ik} \tag{2.20}$$

Где V_W – оптимальной чистая стратегия по критерию Вальда, i, k – выбранные максиминные элементы, нижняя чистая цена игры g - максимальный выигрыш. Выбор такой стратегии подходит очень осторожному игроку.

Как и критерий Вальда, критерий Сэвиджа [25] является критерием крайнего пессимизма - это критерий минимаксного риска. В этом случае ЛПР исходит из предположения, что природа очень злонамеренна и реализует самые неблагоприятные для него состояния. Критерий Сэвиджа рекомендует выбирать в качестве оптимальной ту чистую стратегию S_i , при которой минимизируется величина максимального риска.

В отличие от критерия Вальда, данный критерий базируется не на матрице выигрышей, а на матрице риска, и в нашем случае выглядит как

$$V_S = \min_i \max_k r_{ik} \tag{2.21}$$

Наиболее предпочтительным является критерий Гурвица [23] - это критерий пессимизма – оптимизма, или его можно назвать критерием здравомыслия. При применении этого критерия рекомендуется руководствоваться рациональной оценкой возможного состояния природы и выигрышных шансов игрока:

$$V_H = \max_i \{\theta^{\circ} * \min_k g_{ik} + (1 - \theta^{\circ}) \max_k g_{ik} \tag{2.22}$$

Применяемый здесь коэффициент θ°- это так называемый коэффициент оптимизма, и он выбирается на основании субъективных соображений игрока.

Чем больше желание подстраховаться в данной ситуации, тем ближе к единице значение. При θ°=0 получается критерий максимакса или крайнего оптимизма; при θ°=1 - критерий крайнего пессимизма Вальда.

Поскольку природа не будет постоянно в одном чистом состоянии, то можно утверждать, что природа применяет смешанные стратегии в соответствии с какой-то закономерностью или случайным образом. В соответствии с этим наш игрок будет также использовать смешанные стратегии с определенными вероятностями:

$$p = (p_1, p_2, \dots, p_m)^T \tag{2.23}$$

При этом

$$p = V * q \tag{2.24}$$

Рассмотрев две игры, мы получим, естественно, два разных решения, что нашего игрока никоим образом не устраивает, поскольку, выиграв в одной игре, можно проиграть в другой. Необходимо теперь найти компромиссное решение, дающее приемлемый выигрыш по результатам игр с двумя соперниками.

Предпринимательским структурам приходится осуществлять свою деятельность в условиях риска и неопределенности, все чаще руководства компаний обращается к математическим методам и моделям, что объясняется возможностью формализации, конструктивного развития и повышения эффективности методов стратегического управления. [26] Для повышения эффективности системы стратегического управления целесообразно использовать теорию игр, позволяющую найти оптимальное решение с учетом представлений об участниках процессов управления, их ресурсных возможностях и потенциале, а также возможных поступках с учетом существующих рисков. [27]

Игра противника с природой строится на платежной матрице, показанной в таблице 1. Следует отметить, что q вероятности состояния природы остаются теми же, поскольку игра с природой происходит в одних и тех же временных рамках, а g - возможные выигрыши отличаются.

Таблица 1. Платежная матрица игры разумного противника с природой в кризисных условиях

		Состояния природы					
		Q_1	Q_2	…	Q_k	…	Q_l
Стратегии разумных противников	Q_1	g_{11}	g_{12}	…	g_{1k}	…	g_{1l}
	Q_2	g_{21}	g_{22}	…	g_{2k}	…	g_{2l}
	…	…	…	…	…	…	…
	Q_j	g_{j1}	g_{j2}	…	g_{jk}	…	g_{jl}
	…	…	…	…	…	…	…
	Q_n	g_{n1}	g_{n2}	…	g_{nk}	…	g_{nl}
Вероятности состояния		q_1	q_2	…	q_k	…	q_l

Определено, что для повышения эффективности системы стратегического управления целесообразно использовать теорию игр, позволяющую найти оптимальное решение с учетом представлений об участниках процессов управления, их ресурсных возможностях и потенциале, а также возможных поступках с учетом существующих рисков. Разработана платежная матрица игры разумного противника с природой в кризисных условиях.

Математической моделью задачи принятия решений выступает статистическая игра, т.е. игра с природой. В данном случае, в отличие от парных матричных игр, сознательно действует только один из игроков, чаще всего называемый активным, который и выступает в качестве ЛПР. Второй игрок - пассивный игрок или «Природа», который представляет собой внешнюю среду, влияющую на результат принятого ЛПР решения - не принимает решений и фактически является абсолютно нейтральным как к выигрышу, так и к проигрышу ЛПР. При использовании статистических игр для моделирования принятия решений на практике в роли ЛПР может выступать один человек (менеджер, директор, начальник отдела и т.д.) или группа лиц (совет директоров, собрание акционеров и т.д.). Можно отметить, что статистические игры достаточно наглядны для иллюстрации принятия решения в вопросах инвестирования, так как рыночная конъюнктура не играет против инвестора, а изменяется под действием многих факторов.

Методы принятия решений в статистических играх зависят от характера неопределенности, точнее от того, известны или нет вероятности возможных состояний (стратегий) природы. В случае, когда вероятности наступления того или иного состояния среды известны, неопределенность называется частичной, в обратном же случае - полной неопределенностью [28].

Пусть ЛПР имеет n возможных стратегий поведения: $a_1, a_2, \dots, a_n$; а природа, в свою очередь, может оказаться в одном из m возможных состояний: $g_1, g_2, \dots, g_n$; кроме того, для каждой допустимой комбинации ((a-, g} -) известно значение h_j- количественная оценка эффективности (выигрыш) от использования ЛПР стратегии a - при состоянии природы g. Таким образом, игру можно представить в виде платёжной матрицы, элементами которой будут являться выигрыши от определенных комбинаций стратегий обоих игроков. Для случая частичной неопределенности предполагаются известными вероятности состояний природы $q_j, j = 1, n$. Так как пассивный игрок не имеет никакой заинтересованности в результатах игры, то уменьшение размерности платежной матрицы в статистических играх возможно только за счет наличия доминируемых стратегий у активного игрока.

Так как платежная матрица статистической игры отражает эффективность результата реализации принятого решения, то и при применении указанных выше критериев определение оптимальных стратегий было основано на максимизации итогового значения. Заметим, что критерий крайнего оптимизма и максиминный критерий Вальда связаны между собой через критерий Гурвица. Для использования данного критерия необходимо задать параметр λ (0 ≤ λ ≤ 1). Расчет итогового показателя критерия осуществляется по формуле:

$$\max_i \gamma_i = \max_i(\lambda * \min_j h_{ij} + (1-\lambda) * \max_j h_{ij}) \quad (2.25)$$

и при 1 = 0 критерий Гурвица идентичен критерию крайнего оптимизма, при 1 = 1 - критерию Вальда. Параметр 1 выступает показателем пессимизма ЛПР и отражает его отношение к предполагаемому развитию сценария состояний природы.

При применении критерия Байеса были исследованы три случая, в зависимости от предполагаемого сценария поведения пассивного игрока. Кроме того, использование критерия Ходжа-Лемана позволило продемонстрировать связь критериев Байеса и Вальда. Для использования данного критерия необходимо ввести и задать параметр u $(0<u<1)$. Расчет итогового показателя критерия осуществляется по формуле: [29]

$$\max_i (hl)_i = \max_i(u * \sum_{j=1}^{n} h_{ij} q_j + (1-u) * \min_j h_{ij}) \quad (2.26)$$

и при $u=0$ критерий Ходжа-Лемана идентичен критерию Вальда, при $u=1$ - критерию Байеса. Параметр и выступает показателем степени доверия ЛПР к имеющейся информации о предполагаемых возможных вероятностях состояний природы.

Неопределенность финансового рынка лежит в основе резкого изменения курсов ценных бумаг. Но неопределенность отличается от риска. Риски могут быть сведены к вероятности и в результате определены количественно [30]. Например, шансы дожить до определенного возраста можно рассчитать актуарно. После оценки риск можно застраховать.

Под неопределенностью понимается неспособность предсказать последствия или результаты из-за недостатка знаний или оснований для каких-либо прогнозов. Этот термин часто широко используется в финансовой отчетности, особенно потому, что существует множество событий, находящихся вне контроля компании, которые могут сильно повлиять на ее транзакции. Поскольку в периоды неопределенности принимать финансовые решения намного сложнее, многие владельцы компаний воздерживаются от их принятия, чтобы не создавать проблем.

Событие является неопределенным, если оно имеет неизвестную вероятность. Это иллюстрирует важные последствия различия. Люди могут предпочесть азартные игры с точными вероятностями азартным играм с неизвестными шансами. Такое поведение несовместимо с моделью ожидаемой полезности. Поскольку неопределенность, в отличие от риска, может оказывать значительное влияние на индивидуальное поведение, она также должна быть важным фактором, определяющим равновесные результаты. Неопределенность, вероятно, должна привести к заметным отклонениям от стандартного поведения распределения рисков в ожидаемых моделях полезности.

Путем тщательного анализа - фундаментального или технического - и исследования каждой позиции, которую вы хотите открыть, можно определить тенденции и закономерности среди хаотических рыночных движений. Всегда будет элемент случайного поведения на рынке, но трейдеры могут снизить риск непредсказуемых движений с помощью стратегии управления рисками.

Трейдеры, придерживающиеся теории случайного блуждания, будут считать, что невозможно превзойти фондовый рынок, и попытка сделать это повлечет за собой большой риск. Сторонники этой гипотезы склонны придерживаться стратегии «покупай и держи», поскольку теория предполагает, что долгосрочные позиции будут иметь наибольшие шансы на успех.

Трейдеры будут стремиться иметь разнообразный выбор акций, которые наилучшим образом представляют весь фондовый рынок - торгуемые на бирже фонды (ETF) и индексы являются популярными инструментами, поскольку они отслеживают цены на акции ряда компаний.

Анализ ценных бумаг и управление портфелем - это новое направление, которое применяется для понимания движения любых ценных бумаг, котирующихся на бирже. Хотя классические методы часто используются для оценки и измерения неопределенности, связанной с ценными бумагами и портфелями, эти методы в основном детерминированы, но в реальном мире движение этих ценных бумаг по своей природе является стохастическим поведением.

Для количественной оценки динамической системы применяются инструменты из сложных сетей для изучения последовательных во времени цен фондового рынка [31]. Как правило, наиболее доступные сетевые подходы отображают временные ряды в сетевой домен, чтобы он представлял топологические и структурные свойства системы. Структура сообщества сетей фондового рынка представляет собой структурные изменения во время финансового кризиса.

Изучение сетей фондового рынка не только улучшает решения, связанные с промышленными предприятиями, но также обеспечивает надежный индикатор неизбежного повсеместного снижения стоимости акций, которое относится к финансовому кризису. Это описание эволюции сети имеет тенденцию передавать динамичный финансовый рынок, который предполагает лежащую в основе финансовую деятельность и партнерства

Глава 2. Модель прогнозирования

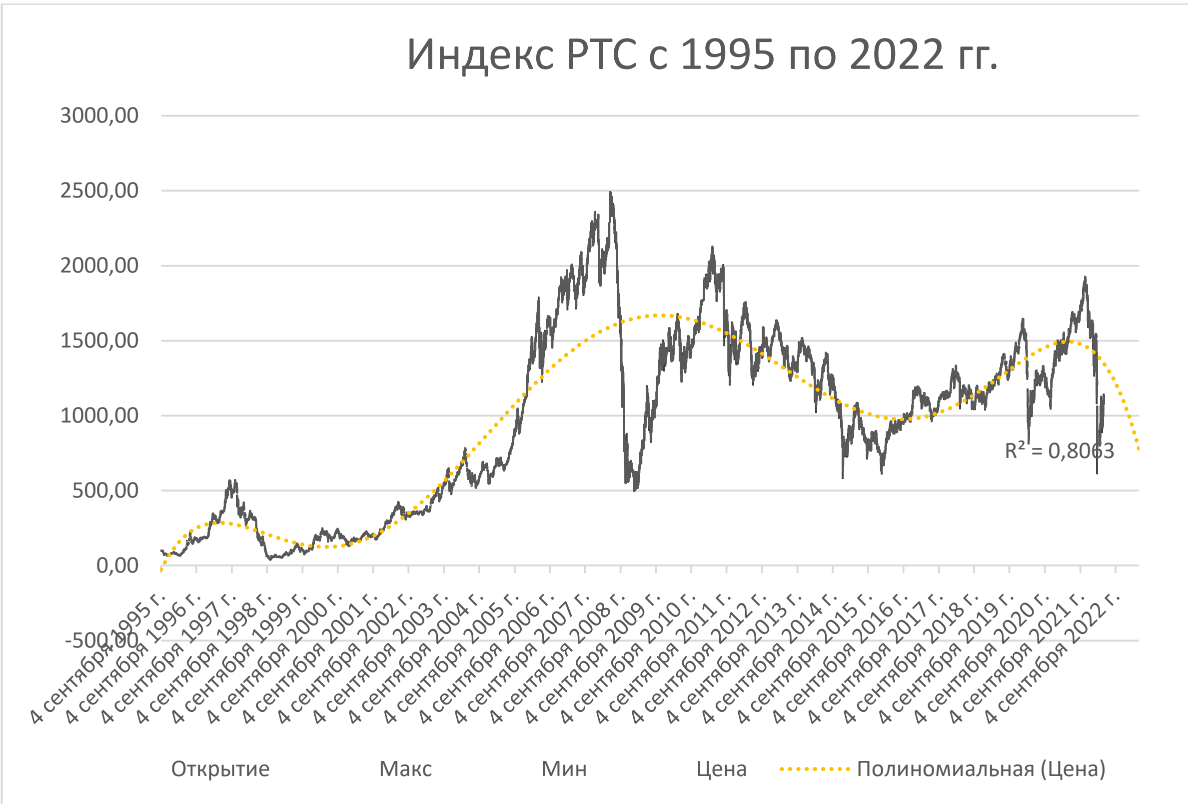
2.1. Составление алгоритма для модели прогнозирования

Исходные данные были взяты из сайта MICEX — Investing.com // Investing.com [32]. Выборка бралась по дням, начиная с 23 Сентября 1997 года для Индекса Московской Биржи, 4 Сентября 1995 года для Индекса РТС и Индекса Голубых Фишек. Данные индекса Московской биржи и индекса РТС представляют собой композитный фондовый индекс, включающий 50 наиболее ликвидных акций крупнейших российских эмитентов. Перечень эмитентов и их вес в индексе пересматривается раз в квартал. Индекс Голубых фишек представляет собой акции 15-ти наиболее ликвидных эмитентов российского фондового рынка. (Приложение 1).

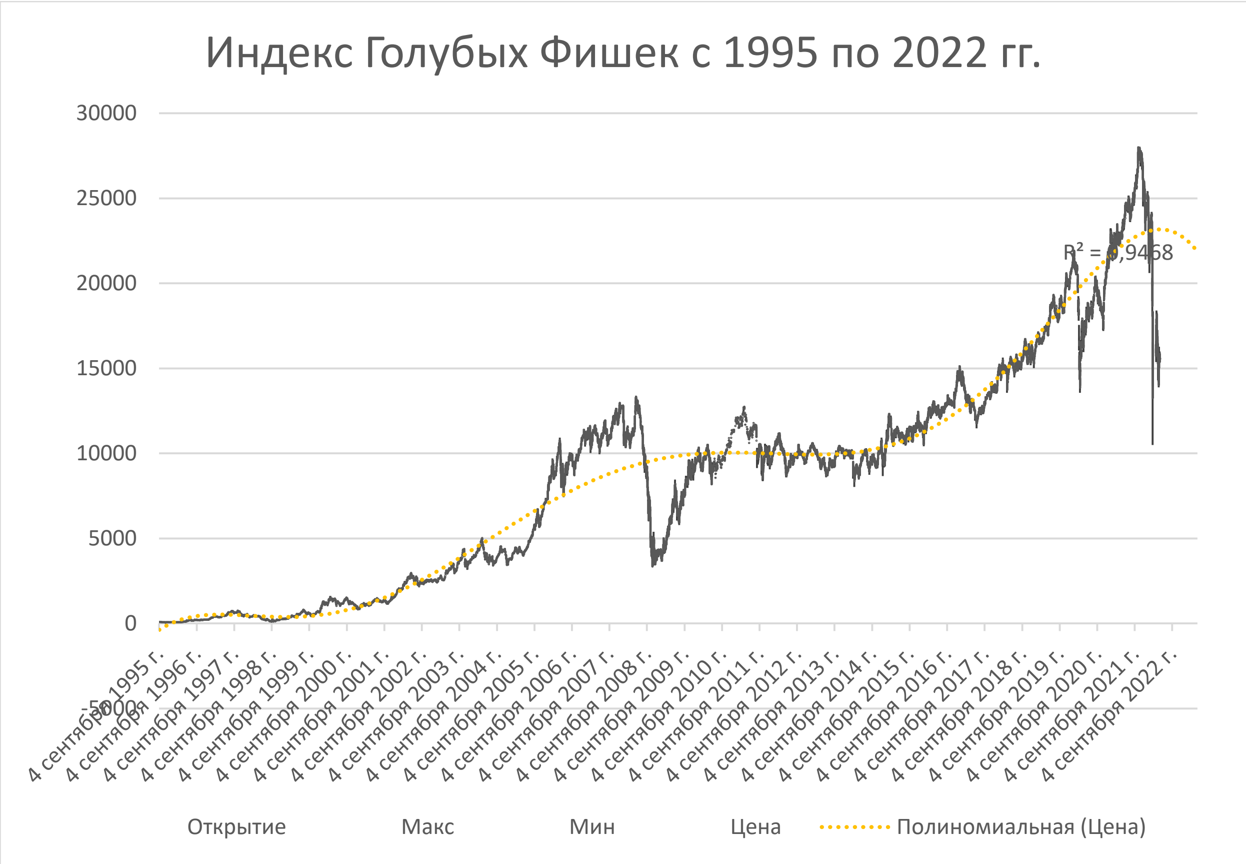
Наглядное отображение данных было сделано с помощью программы Microsoft Excel. В табличном варианте Excel данные за 2022 год содержатся в Приложении 1. В нем указаны, помимо даты, цена, максимальные и минимальные значения, цена на момент открытия.

Данные представлены в виде временных рядов на Рис. 3. С помощью встроенной возможности Excel можно провести полиномиальный тренд с 6 степенью, что повышает величину достоверности аппроксимации (R^2). Также тренд был продолжен до конца 2022 года.

1) Индекс московской биржи



2) Индекс РТС



3) Индекс голубых фишек

Рис. 3. Исходные данные в Excel

Полученные значения R^2 для Индекса Московской Биржи, РТС, Голубых фишек соответственно: 0,9374; 0,8063 и 0,9468. Чем ближе значение R^2 к единице, тем лучше теоретическое распределение описывает реальное.

Условно там же можно вывести уравнения полиномиальных трендов:

y = -4E-19x6 + 9E-14x5 - 9E-09x4 + 0,0005x3 - 13,727x2 + 216158x - 1E+09;

y = -6E-19x6 + 1E-13x5 - 1E-08x4 + 0,0007x3 - 21,033x2 + 332824x - 2E+09;

y = -3E-18x6 + 6E-13x5 - 6E-08x4 + 0,0031x3 - 93,014x2 + 1E+06x - 1E+10.

Наглядно заметно, что корреляция между ними имеет место. Но нас интересует прежде всего составление их прогноза. Для удобства, представим их на едином рис. 4:

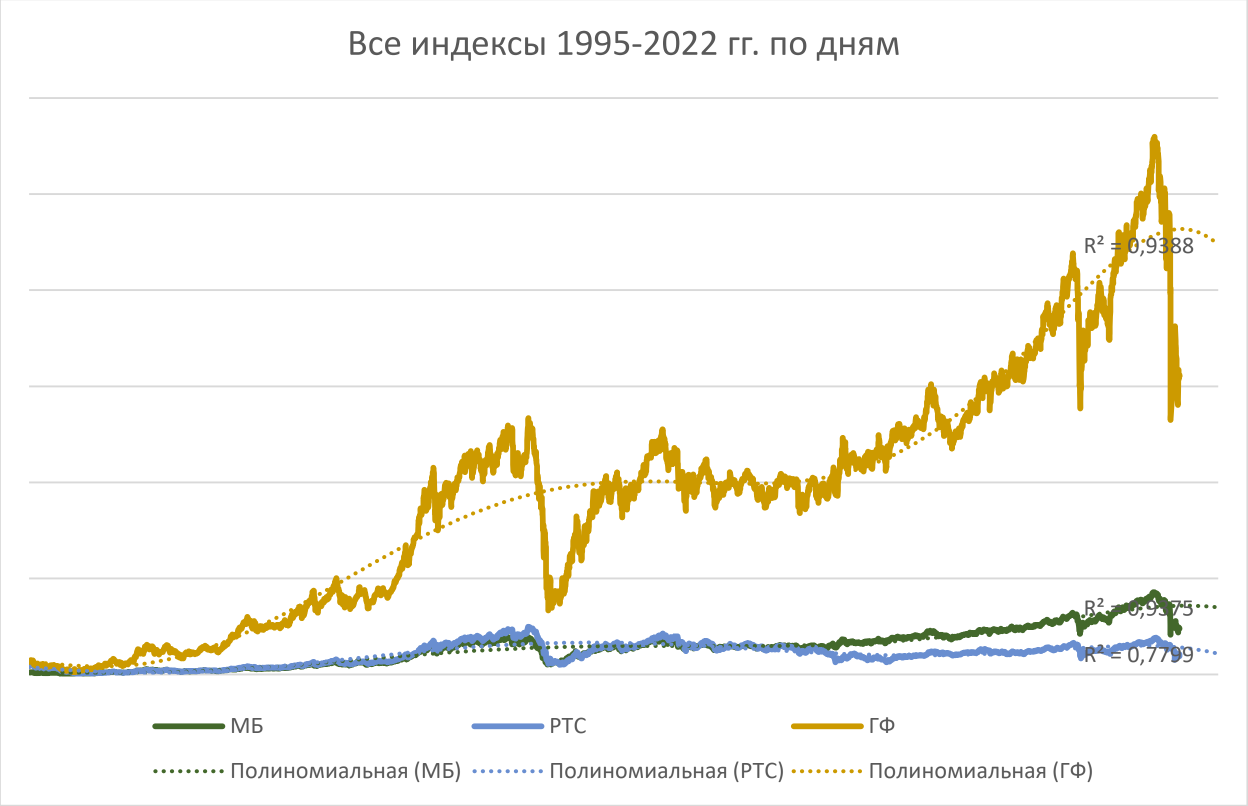


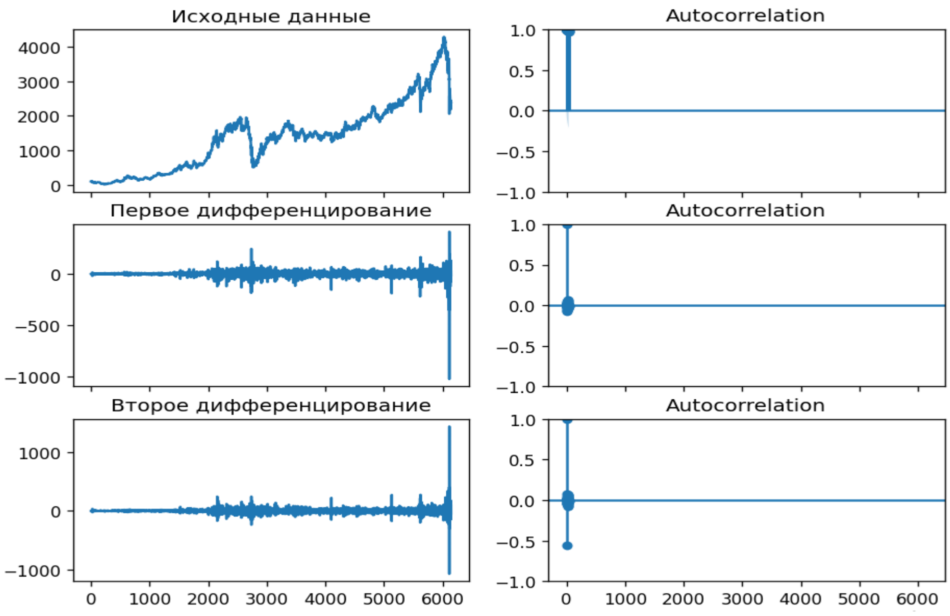
Рис. 4. Индексы котировок, представленные на одном графике.

Ранее в главе 1 было изложено и подробно обосновано, что оптимальным методом для работы с рядами данных такого типа с целью получения прогноза является подход, основанный на методологии Бокса-Дженкинса - модели ARIMA и ее улучшения SARIMAX. Для проведения исследования в рамках данной магистерской диссертации алгоритм указанной модели реализован на языке программирования Python 3.8 в программной среде Jupiter Netbook. Сам код представлен в Приложении 2.

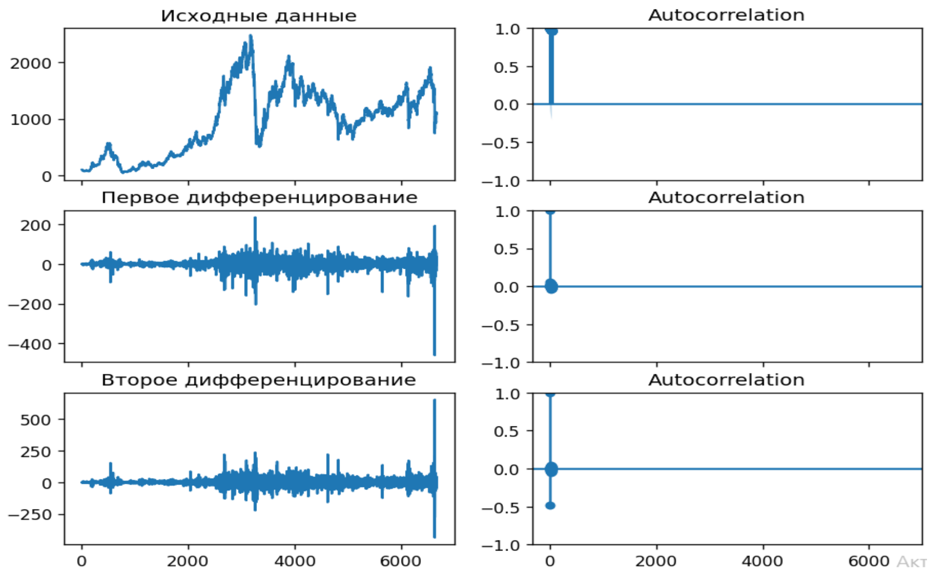
2.2. Проверка исходных данных

В качестве проверки нулевой гипотезы о том, что единичный корень присутствует в авторегрессионной модели заданного временного ряда был применен тест Дики-Фуллера (DF-тест, Dickey-Fuller test) [33] — это методика, которая используется в прикладной статистике и эконометрике для анализа временных рядов на стационарность. Является одним из тестов на единичные корни (Unit root test).

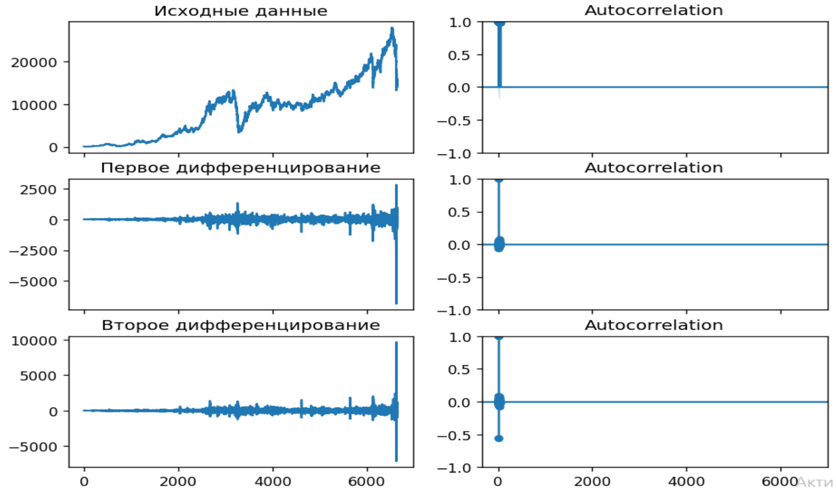
Если в тестовые регрессии добавить лаги первых разностей временного ряда, то распределение DF-статистики (а значит, критические значения) не изменится. Такой тест называют расширенным тестом Дики — Фуллера (Augmented DF, ADF) [34]. Результаты применения теста представлены на рис. 5.



1) Индекс московской биржи



2) Индекс РТС

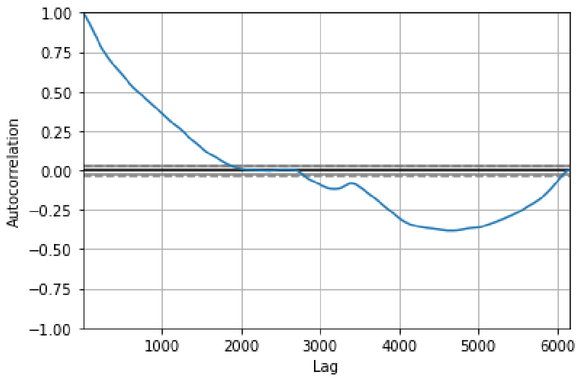


3) Индекс голубых фишек

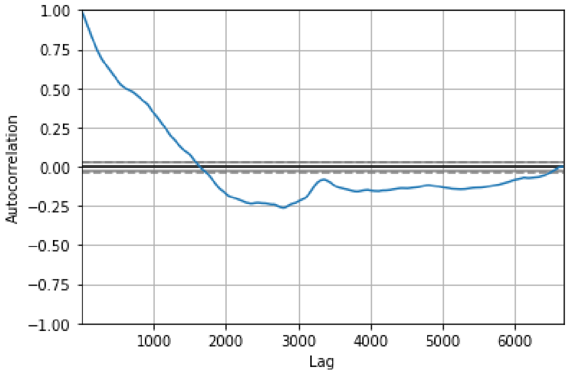
Рис. 5. Дополненный тест Дики-Фулера.

С помощью программы Python были получены следующие результаты: Расширенный тест Дики — Фуллера: -1.306764 для Московской Биржи; -2.093305 для Индекса РТС; -1.202900 для Голубых Фишек.

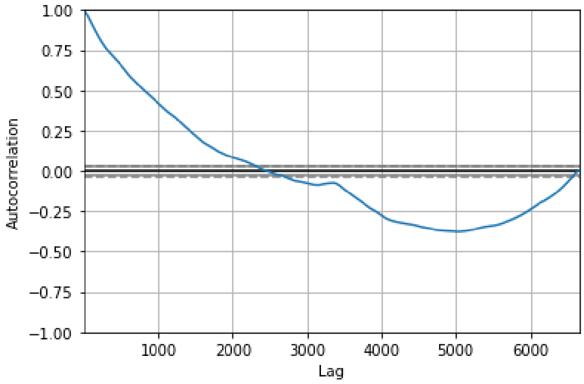
Расширенная статистика Дики – Фуллера (ADF), используемая в тесте, представляет собой отрицательное число. Чем меньше значение, тем очевиднее отклонение гипотезы о существовании единичного корня на некотором уровне значимости. Полученные результаты говорят о том, что никто из них не имеет интегрированный временной ряд и никто из них не стационарен. Автокорреляция отсутствует.



1) Индекс московской биржи



2) Индекс РТС



3) Индекс голубых фишек

Рис. 6. Автокоррецяция данных в зависимости от лага.

На Рис. 6 преставлен график автокореляции исходных данных в зависимости от лага.

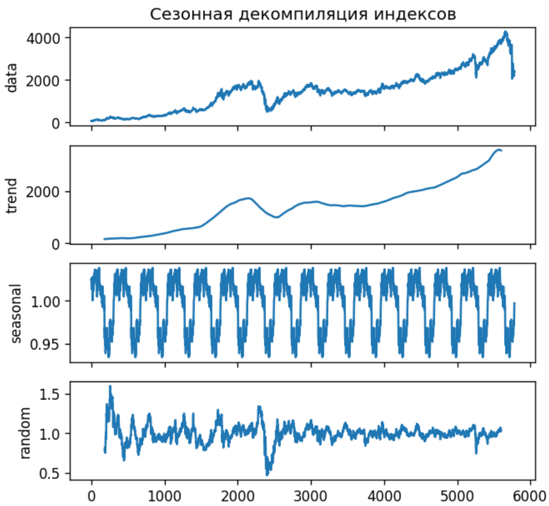
В результате применения теста были получены p - значения для Индекса Московской Биржи, РТС, Голубых фишек соответсвенно: 0.626072, 0.247215, 0.672347.

При тестировании значимости нулевой гипотезы p-значение – это вероятность получения результатов теста, по крайней мере, столь же экстремальных, как и фактически наблюдаемые результаты, в предположении, что нулевая гипотеза верна. P-значение, меньшее, чем уровень значимости, означает, что у такого экстремума наблюдаемый результат был бы очень маловероятен при истинности нулевой гипотезы. [30]

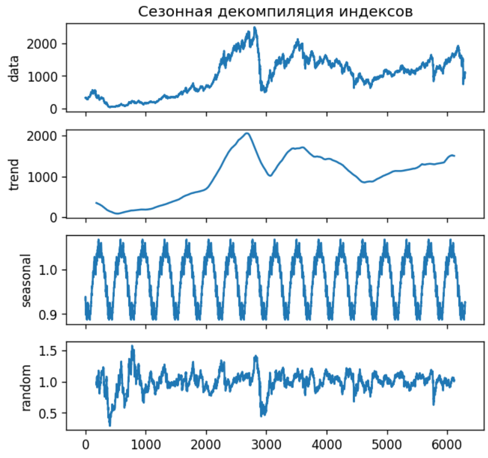
Исходя из вышесказанного, можно заключить, что тесты индексов Московской Биржи и Голубых фишек имеют статистически значимые результаты – они больше, чем уровень знаимости 0,5.

Для того, чтобы провести декомпозицию временного ряда, необходимо определить, является ли она аддитивной, или мультипликативной. По мере увеличения значений временного ряда увеличивается и сезонная изменчивость. Здесь мы должны использовать мультипликативную модель.

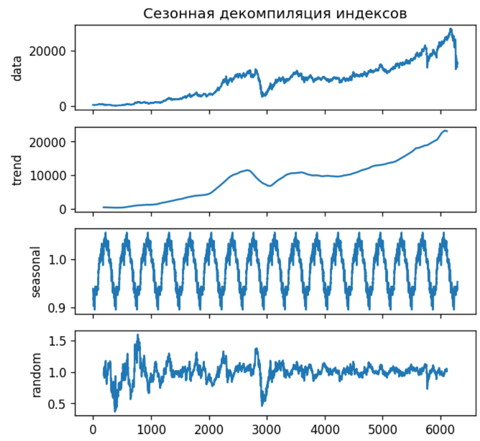
На рис. 7 продемонстрировано разложение исходных рядов на трендовую, сезонную и остаточную оставляющую. Чтобы обнаружить тренд без влияния сезонности и остаточного шума, сглаживаем временные ряды, используя “центрированную скользящую среднюю“. Удаление данного рассчитанного тренда из временного ряда приведет к созданию нового временного ряда, который четко выявляет сезонность. Складываем сезонность вместе и делим на период сезонности. Предыдущие шаги уже извлекли большую часть данных из исходного временного ряда, оставив только “случайный” шум.



1) Индекс московской биржи



2) Индекс РТС



3) Индекс голубых фишек

Рис. 7. Сезонная декомпозиция индексов.

Также проверим данные на наличие единичного корня с помощью альтернативных тестов: KPSS Test, PP Test. При выполнении теста KPSS (тесты Квятковского–Филлипса–Шмидта–Шина) для временного ряда $y_t, t = 1, \ldots, T$ используется метод наименьших квадратов (МНК), чтобы оценить одно из следующих уравнений:

$$y_t = a_0 + \varepsilon_t, y_t = a_0 + \beta t + \varepsilon_t,$$

где a_0 – значение из временного ряда, βt - детерминированный тренд, ε_t - стационарная ошибка.

PP-тест (тест Филлипса-Перрона) используется при нарушении гипотезы о некоррелированности и гомоскедастичности отклонений в тестируемой модели, что позволяет использовать тест для более широкого класса временных рядов. Используется в случаях наличия ярко выраженной сезонности и структурных сдвигов.

Для МБ, РТС, ГФ: KPSS Test = 1, PP Test = 1.

Для трех индексов значения тестов равны 1. Значит у них у всех есть единичный корень и они не стационарны. Обратное, если их значения равны 0. Нулевая гипотеза опровергается.

2.3. Результаты численных экспериментов

Итак, мы проверили исходные временные ряды. Теперь будем строить на их основе прогнозы до конца 2024 года. Данные значение были получены с помощью библиотеки прогнозирования в Python с тренировкой на исходных данных. Тренды представлена в Таблице 2.

Таблица 2. Основные тренды с учётом прогноза с 2021 по 2024 года

Дата	МФ	РТС	ГФ
1 января 2022 г.	1763,255306	2028,479707	11595,15494
1 февраля 2022 г.	1792,090612	2062,999414	11711,74989
1 марта 2022 г.	1820,925918	2097,519121	11828,34483
1 апреля 2022 г.	1849,761224	2132,038827	11944,93978
1 мая 2022 г.	1878,596531	2166,558534	12061,53472
1 июня 2022 г.	1907,431837	2201,078241	12178,12966
1 июля 2022 г.	1936,267143	2235,597948	12294,72461
1 августа 2022 г.	1965,102449	2270,117655	12411,31955
1 сентября 2022 г.	1993,937755	2304,637362	12527,91449
1 октября 2022 г.	2022,773061	2339,157069	12644,50944
1 ноября 2022 г.	2051,608367	2373,676776	12761,10438
1 декабря 2022 г.	2080,443673	2408,196482	12877,69933
1 января 2023 г.	2109,278979	2442,716189	12994,29427
1 февраля 2023 г.	2138,114285	2477,235896	13110,88921
1 марта 2023 г.	2166,949592	2511,755603	13227,48416
1 апреля 2023 г.	2195,784898	2546,27531	13344,0791
1 мая 2023 г.	2224,620204	2580,795017	13460,67404
1 июня 2023 г.	2253,45551	2615,314724	13577,26899
1 июля 2023 г.	2282,290816	2649,83443	13693,86393
1 августа 2023 г.	2311,126122	2684,354137	13810,45888
1 сентября 2023 г.	2339,961428	2718,873844	13927,05382
1 октября 2023 г.	2368,796734	2753,393551	14043,64876
1 ноября 2023 г.	2397,63204	2787,913258	14160,24371
1 декабря 2023 г.	2426,467346	2822,432965	14276,83865

1 января 2024 г.	2455,302653	2856,952672	14393,4336
1 февраля 2024 г.	2484,137959	2891,472379	14510,02854
1 марта 2024 г.	2512,973265	2925,992085	14626,62348
1 апреля 2024 г.	2541,808571	2960,511792	14743,21843
1 мая 2024 г.	2570,643877	2995,031499	14859,81337
1 июня 2024 г.	2599,479183	3029,551206	14976,40831
1 июля 2024 г.	2628,314489	3064,070913	15093,00326
1 августа 2024 г.	2657,149795	3098,59062	15209,5982
1 сентября 2024 г.	2685,985101	3133,110327	15326,19315
1 октября 2024 г.	2714,820407	3167,630033	15442,78809
1 ноября 2024 г.	2743,655714	3202,14974	15559,38303
1 декабря 2024 г.	2772,49102	3236,669447	15675,97798

Таблица 3. Прогнозные значения Индексов до конца 2024 года.

Дата	МБ	РТС	ГФ
Январь 2022	1677,02	1919.89	11073,54
Февраль 2022	1759,44	2071.8	11629,43
Март 2022	1874,73	2223.06	12359,45
Апрель 2022	1850,64	2220.11	12156,74
Май 2022	1888,86	2290.51	12690,62
Июнь 2022	1574,33	1906.97	10496,01
Июль 2022	1660,42	2063.94	11192,66
Август 2022	1628,43	2053.93	10861,12
Сентябрь 2022	1667,35	2122.5	11286,48
Октябрь 2022	1925,24	2459.88	13090,91
Ноябрь 2022	1753,67	2303.34	12149,79
Декабрь 2022	1495,33	1966.68	10368,79
Январь 2023	1348,92	1646.14	9086,68
Февраль 2023	1027,66	1211.84	6879,83
Март 2023	731,96	773.37	4616,05
Апрель 2023	611,32	658.14	4058,3
Май 2023	619,53	631.89	4174,77
Июнь 2023	624,9	535.04	4173,08
Июль 2023	666,05	544.58	4374,56
Август 2023	772,93	689.63	5274,72
Сентябрь 2023	920,35	832.87	6264,7
Октябрь 2023	1123,38	1087.59	7497,44
Ноябрь 2023	971,55	987.02	6718,52
Декабрь 2023	1053,3	1017.47	7213,29
Январь 2024	1091,98	1066.53	7544,74
Февраль 2024	1197,2	1254.52	8350,36
Март 2024	1237,18	1348.54	8648,04
Апрель 2024	1284,95	1377.65	8876,82
Май 2024	1370,01	1444.61	9568,83

Июнь 2024	1419,42	1473.81	9827,37
Июль 2024	1332,64	1410.85	9154,18
Август 2024	1450,15	1572.48	9955,03
Сентябрь 2024	1436,04	1572.84	9811,96
Октябрь 2024	1332,62	1384.59	9150,58
Ноябрь 2024	1309,31	1339.35	8979,23
Декабрь 2024	1397,12	1479.73	9648,34

Точками на графике представлены данные, полученные и представленные в Таблице 3. Эти данные представляют собой полученный прогноз модели в Python.

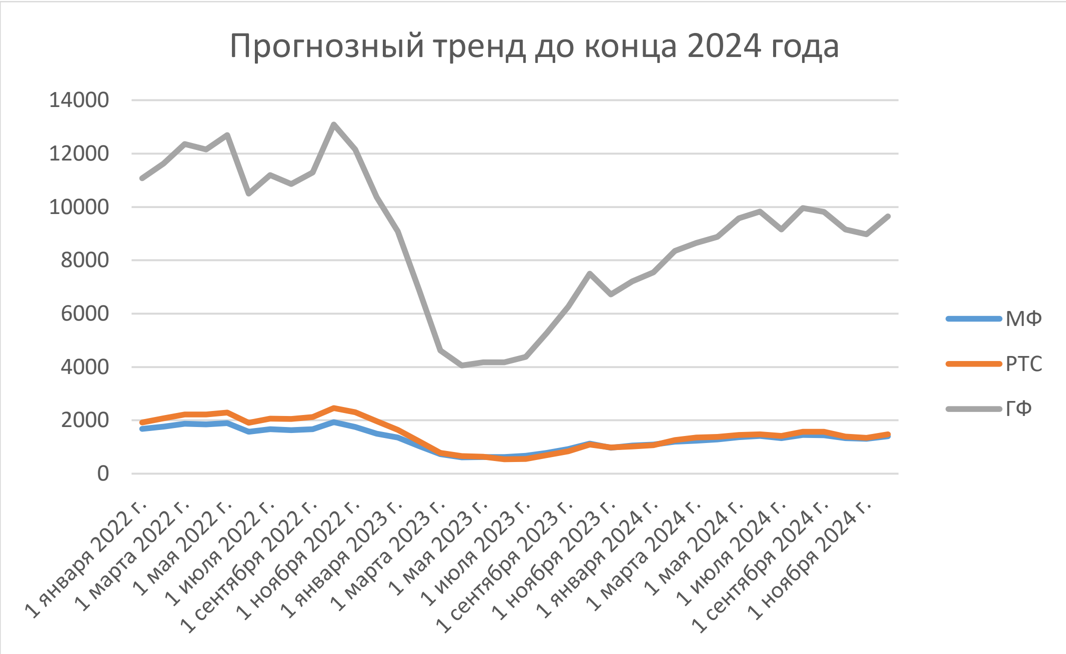


Рис. 8. Прогнозный тренд до конца 2024 года.

Полученные результаты графически представлены на рис. 8. Как эмпирически видно, в мае 2022 года ожидается первый из обвалов, второй – в октябре этого же года.

Рис. 9. Отклонение фактических значений от прогноза.

На рис. 9 сверху изображено наглядное представление отклонения тренда от уже существующих данных за 2022 год. Видно, что, вследствие событий, шкала индекса резко спустился вниз, тем самым ухудшив прогноз. Таким образом, результаты будут иметь плохую значимость в долгосрочном (более года) варианте. Тем не менее, фактические значения остаются в допустимой области, что позволит использовать модель, добавив вероятности доверия.

Вследствие обновления библиотеки ARIMA, теперь вместо него появляется SARIMAX. Это даже лучше, потому что, если данные не являются стационарными, мы можем использовать модель SARIMAX. ARIMA включает в себя авторегрессионную интегрированную скользящую среднюю, в то время как SARIMAX включает сезонные эффекты и экзогенные факторы с авторегрессионной и скользящей средней составляющей в модели. Модель SARMAX позволяет также, как ARIMA строить прогнозы. Результаты применения тестов, содержащихся в данной библиотеке, представлены на Рис. 10.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   Цена   No. Observations:                 6146
Model:                 ARIMA(1, 1, 2)   Log Likelihood              -29215.371
Date:                Thu, 12 May 2022   AIC                          58438.743
Time:                        13:54:10   BIC                          58465.636
Sample:                             0   HQIC                         58448.071
                               - 6146
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.0597      0.127      0.471      0.638      -0.189       0.308
ma.L1         -0.1270      0.128     -0.989      0.323      -0.379       0.125
ma.L2          0.0405      0.008      4.870      0.000       0.024       0.057
sigma2       789.9501      2.333    338.614      0.000     785.378     794.523
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):          21315276.65
Prob(Q):                              0.98   Prob(JB):                         0.00
Heteroskedasticity (H):              27.59   Skew:                            -8.12
Prob(H) (two-sided):                  0.00   Kurtosis:                       291.07
===================================================================================
```

1) SARIMAX Results для индекса московской биржи

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   Цена   No. Observations:                 6664
Model:                 ARIMA(1, 1, 2)   Log Likelihood              -29967.124
Date:                Thu, 12 May 2022   AIC                          59942.247
Time:                        13:53:16   BIC                          59969.464
Sample:                             0   HQIC                         59951.649
                               - 6664
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.7878      0.104      7.558      0.000       0.584       0.992
ma.L1         -0.7347      0.105     -7.009      0.000      -0.940      -0.529
ma.L2         -0.0249      0.010     -2.451      0.014      -0.045      -0.005
sigma2       472.0997      1.913    246.769      0.000     468.350     475.849
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):             448122.36
Prob(Q):                              0.98   Prob(JB):                          0.00
Heteroskedasticity (H):               9.02   Skew:                             -1.90
Prob(H) (two-sided):                  0.00   Kurtosis:                         43.00
===================================================================================
```

2) SARIMAX Results для индекса РТС

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   Цена   No. Observations:                 6658
Model:                 ARIMA(1, 1, 2)   Log Likelihood              -44087.719
Date:                Thu, 12 May 2022   AIC                          88183.437
Time:                        13:55:13   BIC                          88210.651
Sample:                             0   HQIC                         88192.838
                               - 6658
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.5874      0.041    -14.315      0.000      -0.668      -0.507
ma.L1          0.5280      0.041     12.748      0.000       0.447       0.609
ma.L2          0.0150      0.006      2.370      0.018       0.003       0.027
sigma2      3.315e+04    109.039    303.997      0.000    3.29e+04    3.34e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.00   Jarque-Bera (JB):           24908323.17
Prob(Q):                              1.00   Prob(JB):                          0.00
Heteroskedasticity (H):              46.30   Skew:                             -8.25
Prob(H) (two-sided):                  0.00   Kurtosis:                        302.21
===================================================================================
```

3) SARIMAX Results для индекса голубых фишек

Рис. 10. Результаты моделирования с помощью модулей библиотки SARIMAX.

Посмотрим данные, полученные в результате проведения модели SARIMAX. По результатам тестов на остатки, можно сказать, что данная модель достаточно точная. По тесту Харке — Бера (JB) видно, что она имеет очень большое значение от нуля. Он проверяет на нормальность распределения ошибок, и чем он больше, тем больше оснований опровергнуть нулевую гипотезу. По тесту Юнга–Бокса (Q) можно определеть, что данные распределены независимо друг от друга. Гетероскедастичность (H), то есть неоднородность, имеет место быть.

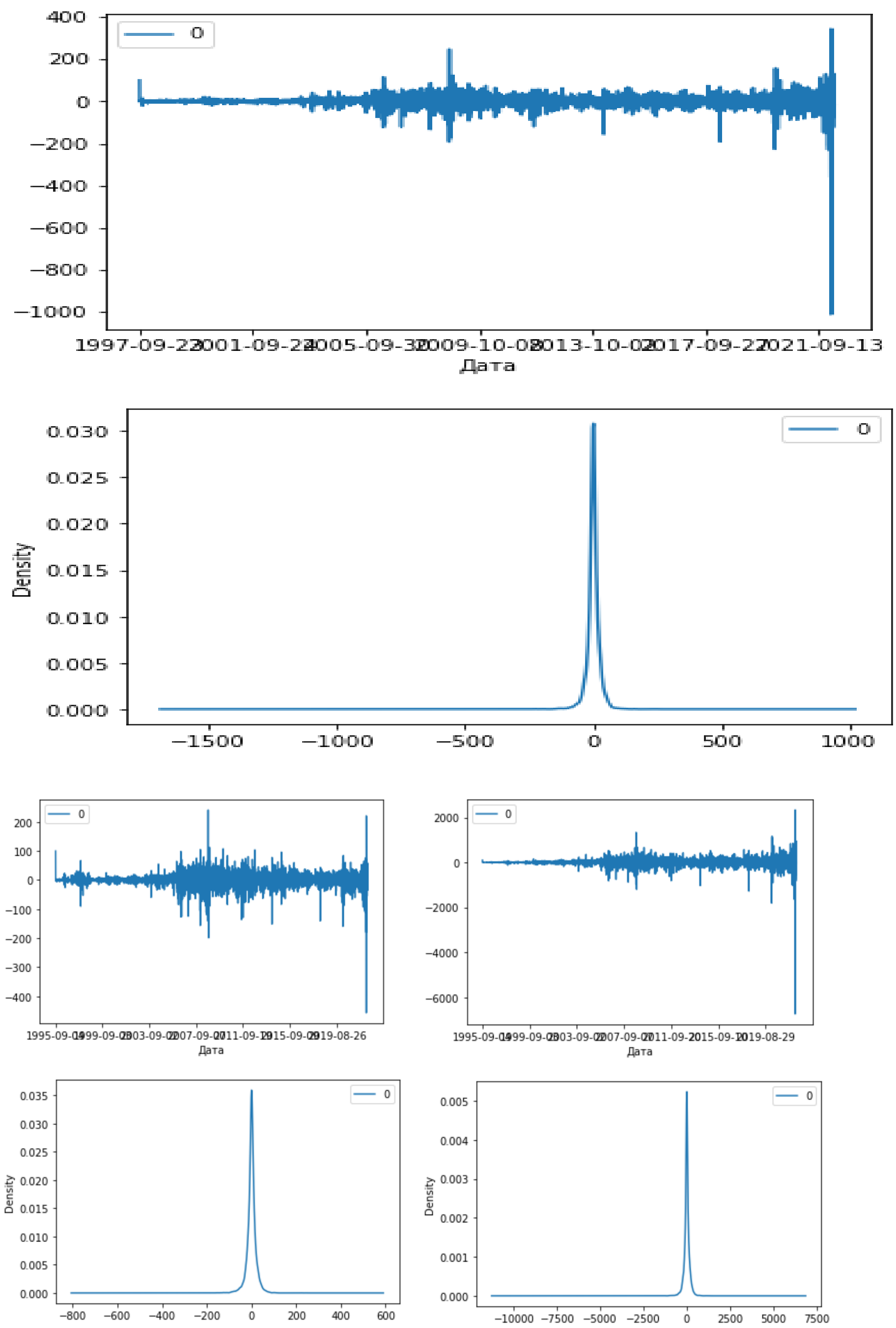


Рис. 11. График плотности ошибок

Во-первых, мы получаем линейный график остаточных ошибок, предполагая, что все еще может быть некоторая информация о тренде, не уловленная моделью.

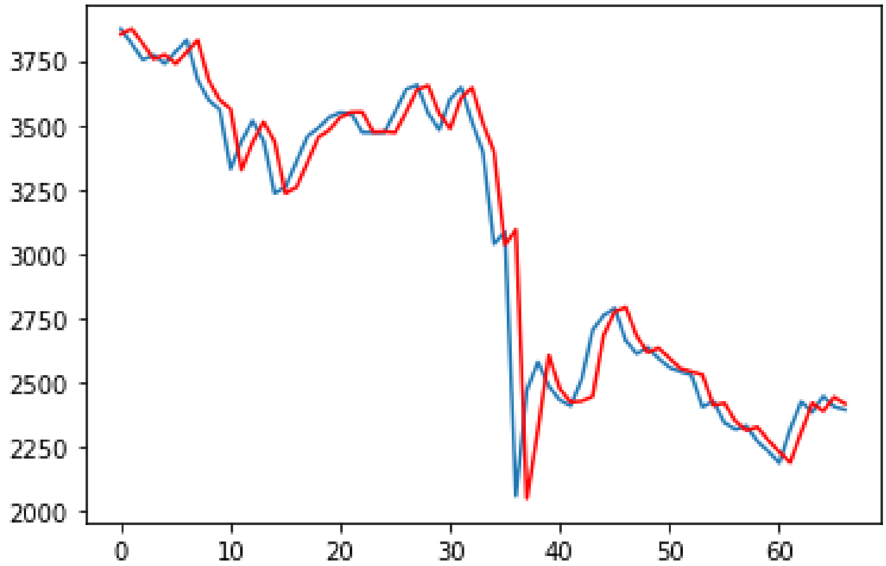
Затем мы получаем график плотности значений остаточных ошибок, предполагая, что ошибки являются гауссовыми, но могут не центрироваться на нуле.

Распределение остаточных ошибок отображается. Результаты показывают, что в прогнозе действительно есть смещение (ненулевое среднее в невязках).

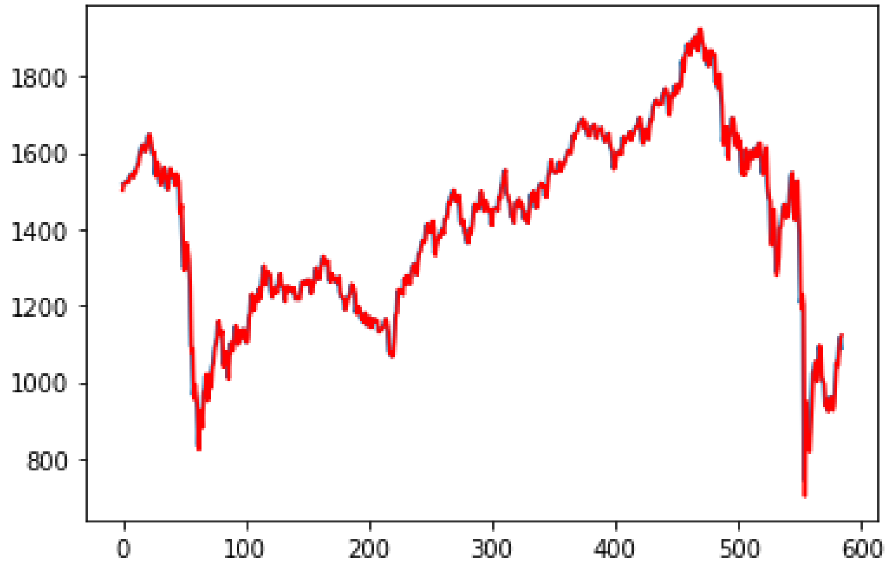
Таблица 4. Описательная статистика индексов

	МБ	РТС	ГФ
Счет	6146	6664	6658
Средняя	0.400	0.146	2.403
Отклонение	28.1159	21.7619	181.9545
Минимум	-1011.237226	-457.347241	-6745.157503
25%	-7.545597	-7.099834	-44.887447
50%	0.639579	0.615814	2.197322
75%	9.642630	8.610231	58.741438
Максимум	344.089169	241.034470	2337.088379

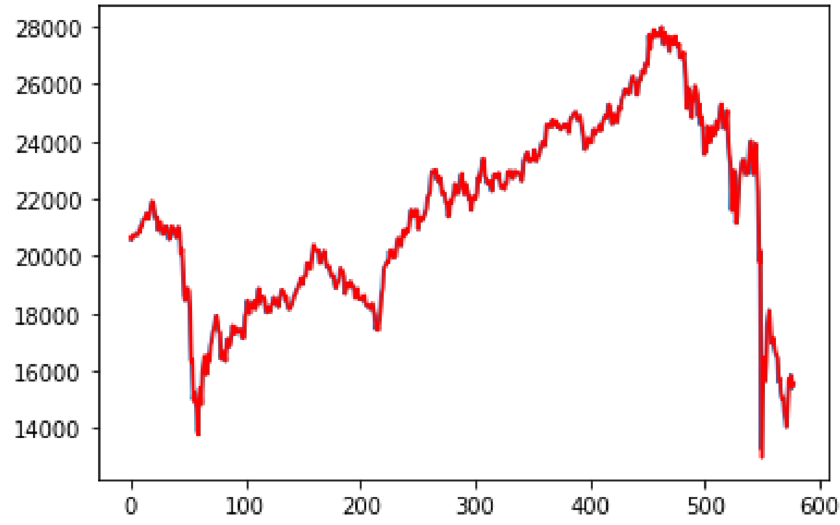
Сравним полученные прогнозы SARIMAX с уже существующими фактическими данными за 2022 год. Синим цветом обозначены фактические данные за 2022 год, а красным – прогноз с с применением SARIMAX. Для индексов РТС и голубых фишек результаты двух выборок полностью совпали и наложились друг на друга.



1) Индекс московской биржи



2) Индекс РТС



3) Индекс голубых фишек

Рис. 12. Прогнозы SARIMAX в сравнении с фактическими данными.

На рис. 12 представлено графическое отображение прогнозов SARIMAX в сравнении с фактическими данными. Видно, что индекс РТС и Голубых Фишек спрогнозованы лучше, чем индекс Московской Биржи. Потому что у ее имеется запаздывание прогноза. Такое запаздывание может быть объяснено меньшей выборкой.

Требуется проверить полученные прогнозы, для этого используем MSE и RMSE. MSE в основном измеряет среднеквадратичную ошибку прогнозов.

Для каждой точки вычисляется квадрат разности между прогнозами и целью, а затем усредняются эти значения. Чем выше это значение, тем хуже модель.

Test MSE: 28105.307 Московская Биржа

Test MSE: 1324.754 РТС

Test MSE: 220433.752 Голубые Фишки

Вторым коэффициентом будет выступать среднеквадратическая ошибка RMSE - это мера того, насколько хорошо линия регрессии соответствует точкам данных. RMSE также можно рассматривать как стандартное отклонение остатков. Индекс Голубых фишек имеет слишком большой разброс остатков. Но это объясняется более высокими показателями.

Тест RMSE Индекса Московой Биржи: 1026.113;

Тест RMSE Индекс РТС: 471.426;

Тест RMSE Индекса Голубых Фишек: 6747.240;

Видно, что РТС имеет наименьшие показатели RMSE (среднеквадратической ошибки), чем остальные.

Итак, судя по результатам тестов на остатки в модели SARIMAX, и проверке на MSE и RMSE, можно утверждать, во-первых, что прогноз построен статистически значимо, для индексов московской биржи и голубых фишек. Во-вторых, по статистике индексов видно, что отклонение невелико.

Прогноз является качественным, потому что лучше всего описывает фактические значения, которые уже имеются с начала 2022 года.

Данная модель частично эффективна, не считая индекса московской биржи, и может быть применен в дальнейших вычислениях.

Глава 3. Поиск оптимальной стратегии

3.1. Модель принятия решения

В этой главе разберем стратегии, которые могут быть приняты игроком при условиях сложившейся ситуации на фондовом рынке, рассматривая выбранные ранее индексы. Поставим перед собой математическую задачу: есть две стратегии ЛПР: купить или продать. Ему нужно принять решение – выбрать одну из них, в зависимости от направления трендов у исследуемых индексов.

Вначале мы получим значения доверителного интервала трендов у интервалов. По столбцам указаны даты, пессимичные и оптимистичные прогнозы индексов московской биржи (МБ), РТС и голубых фишек (ГБ). На второй строчке представлена цена за 3 Января 2022 года, от которого мы и отталкиваемся. Из прогнозов вычитаем цеу на начало года, и получаем выигрыш или проигрыш. Данные представлены на таблице:

Таблица 5. Пессимистичные и оптимистичные прогнозы значений индексов.

Дата	МБ Пес	МБ опт	РТС Пес	РТС Опт	ГФ Пес	ГФ опт
03.01.2022	3853,117	-	1625,012	-	25196,62	-
04.01.2022	3808,947	3897,28785	1585,18	1664,844	24908,97	25484,28
05.01.2022	3791,268	3916,201179	1566,463	1683,94	24789,83	25610,9
06.01.2022	3777,846	3930,857512	1552,155	1698,28	24699,79	25708,42
07.01.2022	3766,628	3943,310261	1540,201	1710,237	24624,64	25791,05
10.01.2022	3756,818	3954,354798	1529,73	1720,708	24558,96	25864,21
11.01.2022	3748,008	3964,39904	1520,298	1730,14	24499,98	25930,66
12.01.2022	3739,957	3973,685305	1511,647	1738,79	24446,08	25992,05
13.01.2022	3732,505	3982,37148	1503,611	1746,827	24396,19	26049,42
14.01.2022	3725,544	3990,567232	1496,073	1754,365	24349,57	26103,52
17.01.2022	3718,993	3998,352409	1488,951	1761,487	24305,68	26154,88
18.01.2022	3712,793	4005,78729	1482,184	1768,254	24264,14	26203,91
19.01.2022	3706,896	4012,918706	1475,723	1774,715	24224,61	26250,91
20.01.2022	3701,265	4019,783899	1469,529	1780,909	24186,86	26296,14
21.01.2022	3695,871	4026,413063	1463,573	1786,865	24150,68	26339,8
24.01.2022	3690,687	4032,831086	1457,829	1792,609	24115,91	26382,05
25.01.2022	3685,694	4039,058765	1452,275	1798,163	24082,4	26423,04
26.01.2022	3680,874	4045,113691	1446,894	1803,544	24050,05	26462,88
27.01.2022	3676,211	4051,010903	1441,671	1808,767	24018,74	26501,67
28.01.2022	3671,693	4056,763376	1436,592	1813,846	23988,39	26539,49
31.01.2022	3667,309	4062,3824	1431,647	1818,791	23958,94	26576,43
01.02.2022	3663,048	4067,877867	1426,825	1823,613	23930,3	26612,55
02.02.2022	3658,902	4073,258504	1422,117	1828,321	23902,43	26647,9
03.02.2022	3654,863	4078,532059	1417,516	1832,922	23875,26	26682,54
04.02.2022	3650,924	4083,705443	1413,015	1837,423	23848,77	26716,51
07.02.2022	3647,079	4088,784858	1408,607	1841,831	23822,9	26749,87
08.02.2022	3643,323	4093,775886	1404,287	1846,151	23797,61	26782,63
09.02.2022	3639,65	4098,683582	1400,05	1850,388	23772,88	26814,84
10.02.2022	3636,055	4103,512535	1395,891	1854,547	23748,68	26846,52
11.02.2022	3632,535	4108,266928	1391,807	1858,631	23724,97	26877,72
14.02.2022	3629,086	4112,950585	1387,792	1862,646	23701,72	26908,44
15.02.2022	3625,704	4117,567015	1383,844	1866,593	23678,93	26938,71
16.02.2022	3622,386	4122,119446	1379,96	1870,478	23656,56	26968,56

17.02.2022	3619,13	4126,610857	1376,137	1874,301	23634,6	26998,01
18.02.2022	3615,931	4131,043999	1372,371	1878,067	23613,02	27027,06
21.02.2022	3612,788	4135,421427	1368,661	1881,777	23591,81	27055,75
22.02.2022	3609,699	4139,745511	1365,003	1885,435	23570,96	27084,08
24.02.2022	3606,66	4144,01846	1361,396	1889,042	23550,44	27112,08
25.02.2022	3603,671	4148,242332	1357,838	1892,6	23530,26	27139,74
24.03.2022	3600,729	4152,419053	1354,327	1896,111	23510,38	27167,1
25.03.2022	3597,832	4156,550426	1350,86	1899,578	23490,81	27194,15
28.03.2022	3594,979	4160,63814	1347,437	1903,001	23471,52	27220,91
29.03.2022	3592,168	4164,683784	1344,055	1906,383	23452,52	27247,4
30.03.2022	3589,397	4168,688851	1340,714	1909,724	23433,79	27273,61
31.03.2022	3586,666	4172,654749	1337,411	1913,027	23415,31	27299,57
01.04.2022	3583,972	4176,582806	1334,146	1916,292	23397,09	27325,27
04.04.2022	3581,316	4180,474276	1330,917	1919,521	23379,11	27350,73
05.04.2022	3578,694	4184,330346	1327,723	1922,715	23361,36	27375,96
06.04.2022	3576,107	4188,152141	1324,563	1925,875	23343,84	27400,95
07.04.2022	3573,553	4191,940726	1321,436	1929,002	23326,55	27425,73
08.04.2022	3571,031	4195,697112	1318,341	1932,097	23309,46	27450,29
11.04.2022	3568,54	4199,42226	1315,277	1935,161	23292,59	27474,65
12.04.2022	3566,08	4203,117087	1312,242	1938,196	23275,91	27498,8
13.04.2022	3563,649	4206,78246	1309,237	1941,201	23259,43	27522,76
14.04.2022	3561,247	4210,419212	1306,26	1944,178	23243,14	27546,53
15.04.2022	3558,873	4214,028132	1303,311	1947,127	23227,04	27570,12
18.04.2022	3556,525	4217,609977	1300,388	1950,05	23211,11	27593,52
19.04.2022	3554,204	4221,165469	1297,492	1952,946	23195,36	27616,75
20.04.2022	3551,909	4224,695298	1294,621	1955,817	23179,78	27639,81
21.04.2022	3549,639	4228,200125	1291,774	1958,664	23164,37	27662,71
22.04.2022	3547,393	4231,680583	1288,952	1961,486	23149,12	27685,44
25.04.2022	3545,171	4235,13728	1286,153	1964,285	23134,02	27708,01
26.04.2022	3542,972	4238,570796	1283,377	1967,061	23119,08	27730,43
27.04.2022	3540,796	4241,981691	1280,624	1969,814	23104,29	27752,7
28.04.2022	3538,641	4245,370501	1277,892	1972,546	23089,65	27774,83
29.04.2022	3536,509	4248,737741	1275,182	1975,256	23075,15	27796,81
04.05.2022	3534,397	4252,083908	1272,492	1977,946	23060,64	27818,79
05.05.2022	3532,306	4255,409479	1269,823	1980,615	23046,14	27840,77
06.05.2022	3530,235	4258,714913	1267,174	1983,264	23031,64	27862,75

Посчитав, получаем данные в процентном выражении, представленные в таблице 6. Эти данные могут быть интерпертированы как таблица выигрышей или прогрышей в зависимости от решения купить или продать ту или иную акцию. Элементами таблицы представлены выигрыши игрока (ЛПР), полученные при такой матричной игре по каждому из индексов по отдельности (московская биржа, РТС, голубые фишки последовательно сверху вниз):

Таблица 6. Выигрыши игрока в зависимости от направления тренда

		Направление тренда	
МБ		Вниз	Вверх
	Купить	-0,91620238	1,105264803
	Продать	0,916202378	-1,1052648
РТС		Вниз	Вверх
	Купить	-0,77979334	1,218830994
	Продать	0,779793345	-1,21883099
ГФ		Вниз	Вверх
	Купить	-0,91406708	1,105822209
	Продать	0,914067076	-1,10582221

Для Московской Биржи моделью принятия решения будет такая матрица выигрышей:

Ожидаемые выигрыши	Направление тренда	
Стратегии	Вниз	Вверх
Купить	-322,88	405,60
Продать	322,88	-405,60

Решением для данной модели представлено на рис. 12. По оси х обозначены стратегии, которые можно принять (1 = купить, 2 = продать). По оси у выражена вероятность принятия того или иного решения. Чем выше данный параметр, тем предподчительней то решение, которое от отражает.

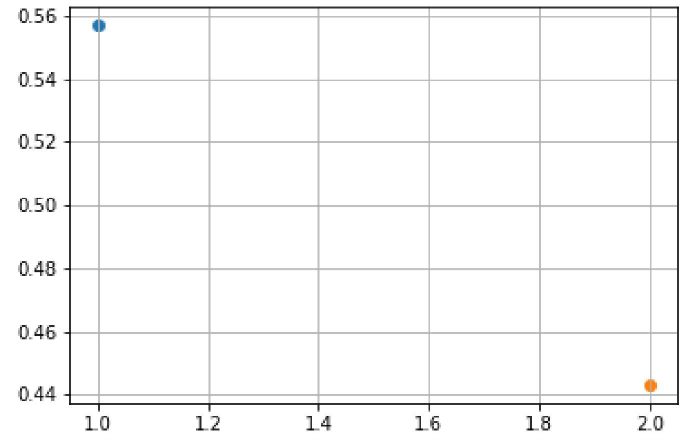


Рис. 13. Принятие решений на онове индекса московской биржи.

Получаем, что первое решение предпочтительнее второго: с вероятностью в 55% покупать выгодней, чем продавать – 45%.

Для Индекса РТС:

Моделью принятия решения будет такая матрица:

Ожидаемые выигрыши	Направление тренда	
Стратегии	Вниз	Вверх
Купить	-357,84	355,60
Продать	357,84	-355,60

Где решением будет:

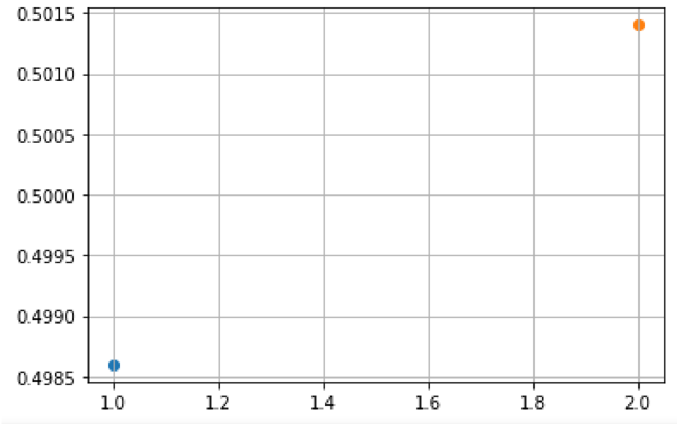


Рис. 14. Принятие решений на основе индекса РТС.

Получаем, что второе решение предпочтительней первого: с вероятностью в 50,14% продавать выгодней, чем покупать – 49,86%.

У Голубых Фишек моделью принятия решения будет такая матрица:

Ожидаемые выигрыши	Направление тренда	
Стратегии	Вниз	Вверх
Купить	-2165,22	2666,36
Продать	2165,22	-2666,36

Где решением будет:

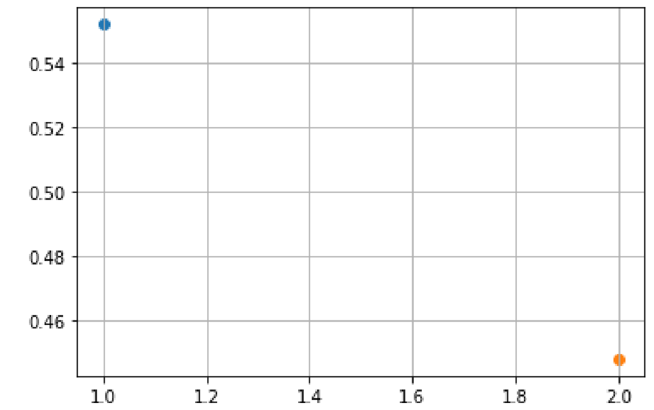


Рис. 15. Принятие решений на основе индекса голубых фишек.

Получаем, что первое решение предпочтительней второго: с вероятностью в 55% покупать выгодней, чем продавать – 45%.

Это значит, что для Московкой биржи и Голубых Фишек предпочтительней первое решение о скупке акций, а для РТС – второе решение – активная продажа.

3.2. Альтернативный способ решения

Рассмотрим альтернативный способ принятия решения с помощью Gym — это среда, созданная компанией OpenAI для тестирования различных моделей обучения с подкреплением, в которой есть как классические модели, так и модели, которые воссоздают физическую среду. В этой среде обучения поддерживается тренировка агента на улучшение результата.

Зададим алгоритму нашей обучающейся программы исходные данные, создав тем самым условие среды, и с ее помощью спрогнозируем и отметим даты, где предпочтительней покупать или продавать:

Пусть случайное начальное значение равно 123 рублю. Средой выполнения будет служить LunarLander-v2. Размер состояния: 32. Размер действия: 2.

LunarLander-v2 – тренажер, в котором в качестве игрока выступает космический корабль. Чем выше он взлетит, тем лучше результат. Как это можно использовать? Во-первых, данная модель использует входящие данные для тренировки, и принимает решения, в зависимости от них. Его задача – максимизировать выигрыш до верхнего значения. Посадочная площадка, то есть, начальная координата всегда находится в (0,0). Координаты - это первые два числа в векторе состояния. Она включает в себя сбор небольшой партии опыта взаимодействия с окружающей средой и использование этой партии для обновления своей политики принятия решений. Как только политика обновляется этой партией, опыт отбрасывается, и новая партия собирается с новой пересмотренной политикой. Вот почему это подход "обучения на основе политики", когда собранные образцы опыта полезны только для обновления текущей политики.

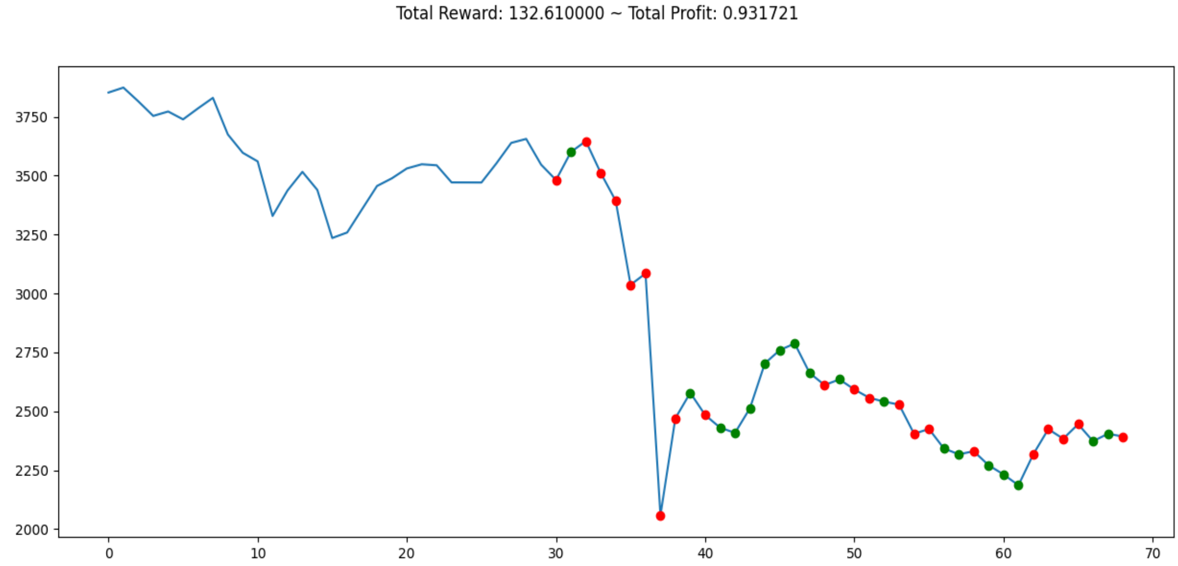


Рис. 16. Индекс МБ с использованием ИИ с подкреплением.

Было проведено сто эпизодов обучения искусственного интеллекта с целью улучшить ее итоговый выигрыш. Получились такие результаты:

На эпизоде 100: средняя длина: 6075, средняя награда: 2989.

Средняя награда за эпизод: 1.3261, средняя длина за эпизод: 0.38.

Общая награда: 132.61.

Общая прибыль: 0.931721.

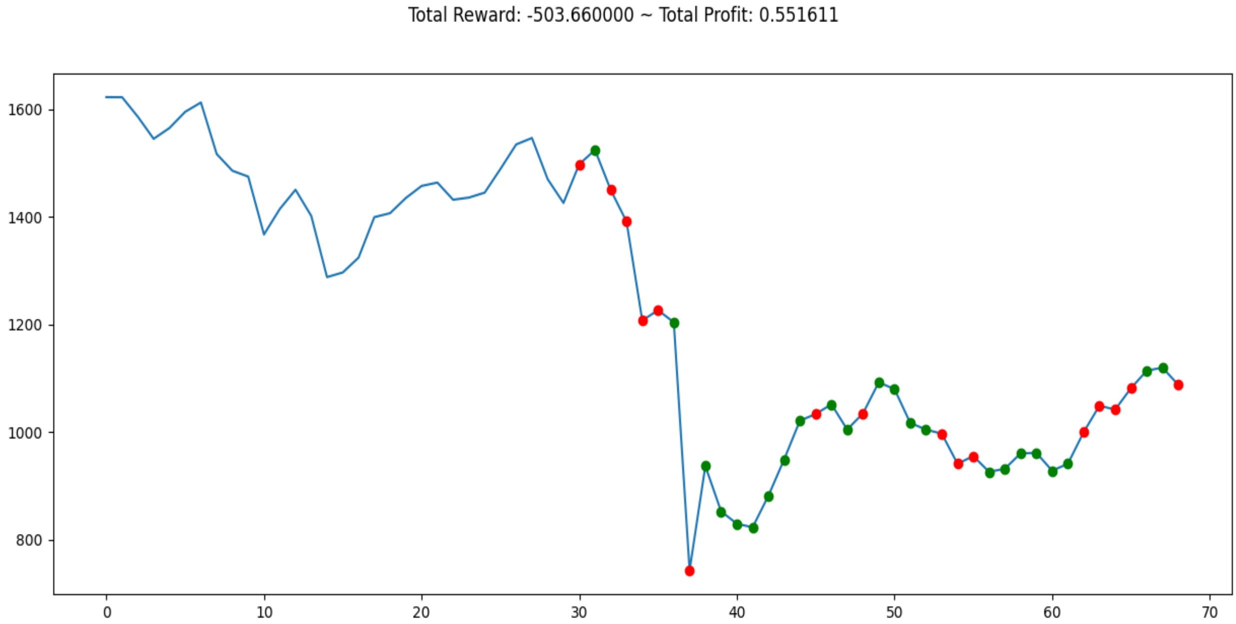


Рис. 17. Индекс РТС с помощью ИИ обучения с подкреплением

На эпизоде 100: средняя длина: 6592, средняя награда: 800.

Средняя награда за эпизод: -5.0366, средняя длина за эпизод: 0.38.

Общая награда: -503.66.

Общая прибыль: 0.551611.

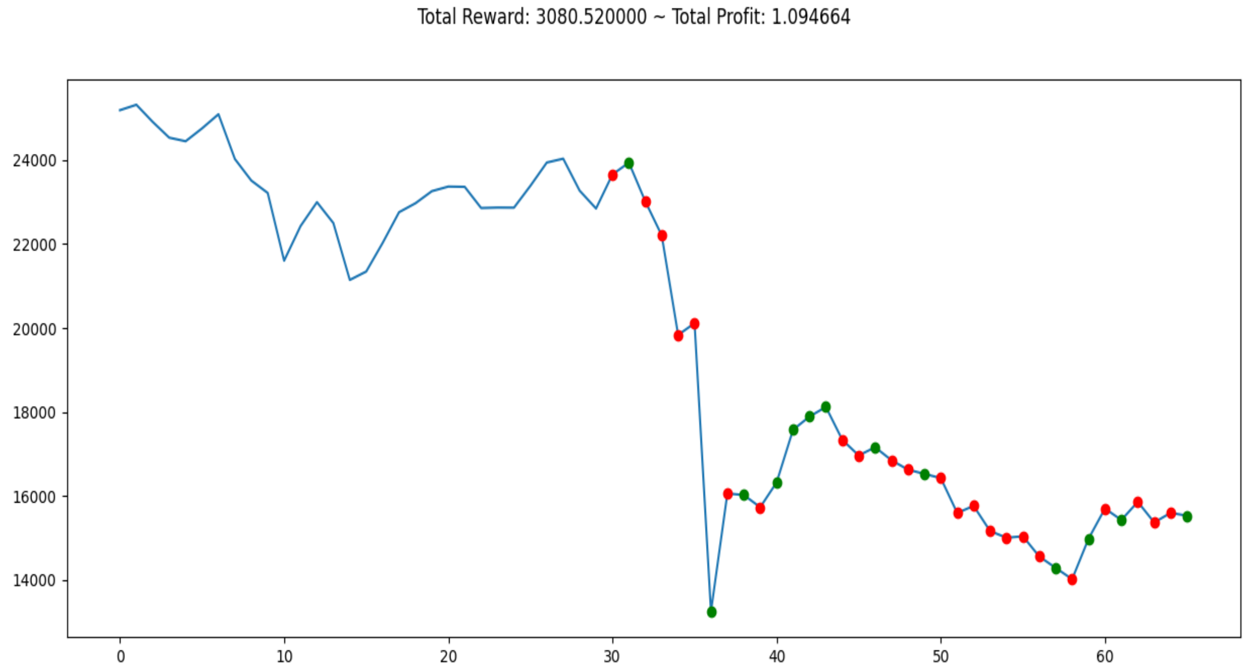


Рис. 18. Индекс Голубых фишек ИИ с обучением

На эпизоде 100: средняя длина:6589, средняя награда: 17095.

Средняя награда за эпизод: 30.8052, средняя длина: за эпизод: 0.35.

Общая награда: 3080.52.

Общая прибыль: 1.09664.

3.3. Полученные результаты и рекомендации

Итак, подведем итоги полученных численных экспериментов. С точки зрения теории принятия решений, предпочтительней покупать акции Индекса Московской биржи и Голубых фишек, но продавать РТС.

Такие решение заданы вероятностно: 55 процентная вероятность успеха стратегии скупки акций индексов московской биржи и голубых фишек, и почти одинаковые вероятности для индекса РТС: 50,14 процента на продажи акций индекса РТС.

С точки зрения искусственного интеллекта все неоднозначно: даже несмотря на рост акций, он его настойчиво продает, получая при этом неплохие результаты.

На эксперименте с индексами московской биржи он заработал 2989 рубля, на РТС – 800 рублей, на голубых фишках – 17095 рублей. В перспективе, можно получить еще больший выигрыш, увеличив количество эпизодов.

Рекомендуется в общем и целом в дальнейшем заниматься покупками акций, входящих в Московскую биржу, но не входящих в Индекс РТС. Представим в виде таблицы 4:

Таблица 7. Рекомендации к скупке-продаже акций на основе индексов

Дата	МБ	РТС	ГФ
янв.22	Покупать	продавать	покупать
фев.22	Покупать	покупать	покупать
мар.22	продавать	продавать	продавать
апр.22	Покупать	покупать	покупать
май.22	продавать	продавать	продавать
июн.22	Покупать	покупать	покупать
июл.22	продавать	продавать	продавать
авг.22	Покупать	покупать	покупать
сен.22	Покупать	покупать	покупать
окт.22	продавать	продавать	продавать
ноя.22	продавать	продавать	продавать
дек.22	продавать	продавать	продавать

Выводы

С помощью программной среды Python был проведен прогноз индексов котировок Московской Биржи. Была проведена проверка на стационарность и определена статистическая значимость моделей. Проведена декомпозиция временных рядов и составлен прогноз до 2024 года. Качество прогноза было проверено с помощью оценки среднеквадратической ошибки.

В ходе работы проведен поиск оптимальной стратегии, с точки зрения полученных результатов прогноза в условиях 2022 года на фондовой бирже, с помощью методологии теории принятия решений. Проводилось составление условий «среды» для решения задачи принятия решений («игры с природой»). Проведен анализ полученных результатов и предоставление рекомендаций по выбору оптимальных стратегий для ЛПР.

В качестве альтернативного способа принятия решения был применен искусственный интеллект, так как он является передовым направлением развития IT-сферы и очень востребован в текущий момент.

Заключение

Оценка перспектив объективно существующих направлений развития общественной деятельности требует формирования прогностического стиля мышления, соответствующих знаний и компетенций, высокой степени общей культуры и образованности, умения видеть и понимать изменения и тенденции окружающей среды, умение использовать научно обоснованные методологии прогнозирования и их инструментарий. Важно понимать, какие перспективные финансовые выгоды можно извлечь из принятия решений или избежать существенных потерь, применяя точные прогнозы.

Неопределенность на финансовых рынках, несомненно, является весьма актуальной проблемой в сегодняшних реалиях. На финансовых рынках одновременно торгуют миллионы различных инвесторов. Каждый участник финансового рынка имеет различный объем собственных средств, обладает разной степенью осведомлённости и придерживается своей стратегии. Это приводит к усложнению структуры рынка, и никто не знает, какие финансовые учреждения купили фиктивные активы и какие учреждения использовали их в качестве обеспечения для получения дополнительных кредитов. Финансовые учреждения все больше рискуют по поводу предоставления кредита, поскольку имеющейся информации недостаточно, чтобы отделить низкий кредитный риск от высокого. На устранение основной неопределенности могут потребоваться месяцы. Но это суть проблематики финансовых рынков, и ее необходимо решать в прямой связи с этими рынками. Задача органов денежно-кредитного регулирования – следить за тем, чтобы проблема не перекинулась на реальную экономику труда и предпринимательства.

Был построен алгоритм поиска оптимальной стратегии в двух вариантах: Python-библиотеки для теории принятия решений и OpenAI_Gum-библиотека с самообучающимся искусственным интеллектом.

Были использованы зависимости поведения субъектов финансового рынка от результатов прогнозных значений для составления матрицы принятия решений. Изучены современные методы статистического прогнозирования. Применена модель Бокса-Дженкинса для прогнозирования данных временного ряда. Проведен процесс нахождения оптимальной стратегии. Весь полученный опыт был апробирован на реальных данных.

В дальнейшем полученные знания могут быть использованы для аналитических и практических методов принятия решений на фондовом рынке, а также в качестве прокторинга экономического состояния в стране.

Список литературы

1. Мировое экономическое положение и перспективы 2021 // Департамент по экономическим и социальным вопросам ООН (DESA). Москва. 2021.
2. Домбровский В.В. Эконометрика: учебник. Москва.
3. Айвазян С.А. Прикладная статистика. Основы эконометрики. 2nd ed. Москва: Юнити-Дана, 2001. 432 pp.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. 6th ed. Москва: Дело, 2004. 576 pp.
5. Бокс Д., Дженкинс Г. Анализ временных рядов, прогноз и управление. 1st ed. Москва: Мир, 1974.
6. Доугерти К. Введение в эконометрику: Пер. с англ. 14th ed. Москва: ИНФРА-М, 1999. 402 pp.
7. Как устроен фондовый рынок [Электронный ресурс] // РБК Инвестиции: [сайт]. URL: https://quote.rbc.ru/news/article/60251b7b9a7947c49c76443d (дата обращения: 22.04.2022).
8. Ханк Д.Э., Уичерн Д.У., Райтс А.Д. Бизнес-прогнозирование. 7th ed. Москва: Вильямс, 2003. 458 pp.
9. Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В. Теория игр: учебник. 2nd ed. СПБ: БХВ-Петербург, 2012. 432 pp.
10. Макаренко А.А. Алгоритмы и программная система классификации полутоновых изображений на основе нейронных сетей, автореферат диссертации на соискание учёной степени кандидата технических наук. Томск. 2017.
11. Садовникова Н.А., Шмойлова Р.А. Анализ временных рядов и прогнозирование. 3rd ed. Москва: ЕАОИ, 2009. 264 pp.
12. Гренджер К., Хатанака М. Спектральный анализ временных рядов в экономике // Статистика. 2012. P. 176.

13. Котенко А.П., Кузнецова О.А. Эконометрика. Временные ряды: метод. указания к лабораторным работам. Самара: Издательство Самарского университета, 2016. 20 pp.
14. Сообщество IT-специалистов [Электронный ресурс] URL: https://radiant-esearpment 88463.herokuapp.eom/ru/eompany/skillfaetory/blog/525214/
15. Terrell D.G., Scott D.W. Variable kernel density estimation // ИНCals of Statistics. 1992. pp. 1236–1265.
16. Суслов В.И., Ибрагимов Н.М. Эконометрия: Учебное пособие. Новосибирск: СО РАН, 2005. 744 pp.
17. Terrell D.G., Scott D.W. Variable kernel density estimation. ИНCals of Statistics, 1992. 1236–1265 pp.
18. Митрофаненко Я.К. Программная система для исследования сезонных трендов физических параметров // Материалы 58-й Междунар. науч.студ. конф. Новосибирск. 2020. P. 240.
19. Суворов А.В. Методы построения макроэкономических сценариев социально-экономического развития // Проблемы прогнозирования. 2016. No. 4. pp. 26-34.
20. Cox J.C., Ross S.A., Rubinstein M. Options pricing: a simplified approach // Journalof Financial Economics. 1979. No. 7. P. 229—263.
21. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов: Учеб. пособие. Москва: Финансы и статистика, 2003. 416 pp.
22. Макаров А.А., Тамбовцева А.А., Василёнок Н.А. Дискретные случайные величины: введение. 2nd ed. Москва: Политология, 2019.
23. Дубров А.М., Лагоша Б.А., Хрусталев Е.Ю. Моделирование рисковых ситуации в экономике и бизнесе: Учеб. пособие. Финансы и статистика, 2000. 176 pp.

24. Лабскер Л.Г., Яновская Е.В. Общая методика конструирования критериев оптимальности решений в условиях риска и неопределенности // Финансовый менеджмент. 2002. No. 5.
25. Паршуков Д.В., Шлепкин А.К., Карпов А.Б. Модели теории игр для выбора оптимальной инновационной стратегии // Вестник КрасГАУ. 2012. No. 5. pp. 116-120.
26. Aguinis H., Henle C.A., Beaty J.C. Virtual Reality Technology: A new tool for personnel selection // International Journal of Selection and Assessment. 2001. No. 9. pp. 70–83.
27. Садовин Н.С., Садовина Т.Н. сновы теории игр. Учебное пособие. Йошкар-Ола. 2011. 119 pp.
28. Грицова О.А., Носырева А.Н., Михайлова О.М. Применение теории игр в управлении персоналом // Региональные проблемы преобразования экономики. 2019. No. 10. pp. 263-268.
29. Чернова Г.В., Кудрявцев А.А. Управление рисками. Проспект, 2005.
30. Пекарский А.В. Математические методы в системном анализе ценовых тенденций фондового рынка. Москва. 2003. 261 pp.
31. Язык программирования Python: сферы применения, методы и этапы изучения [Электронный ресурс] // GeekBrains: [сайт]. URL: https://gb.ru/blog/yazyk-programmirovaniya-python/ (дата обращения: 22.04.2022).
32. ММВБ | Индекс Мосбиржи [Электронный ресурс] // MICEX — Investing.com: [сайт]. URL: https://ru.investing.com/indices/mcx (дата обращения: 27.02.2022).
33. Семенкова Е.В. Управление финансовыми инвестициями: факторы инвестиционной привлекательности // Вестник Российского экономического университета имени Г. В. Плеханова. 2016. No. 2. pp. 77–88.

34. Ni L., Ni Z., Gao Y. Stock trend prediction based on fractal feature selection and support vector machine. 2011. 67 pp.
35. Туртин Д.В., Ершов Б.Л., Капустин Н.А. Поиск оптимальных стратегий на фондовом рынке // Современные наукоемкие технологии. Региональное приложение. 2014. No. 2.
36. Российская экономика: под влиянием кредитного цикла // Министерство экономического развития Российской Федерации, Aug 2019.
37. Энгель, Чарльз М. и Гамильтон, Джеймс Д. Длинные колебания обменного курса: есть ли они в данных и знают ли рынки об этом? Ноябрь 1989.
38. Bremner D., Demaine E., Erickson J., Iacono J., Langerman S., Morin P., Toussaint G. Output-sensitive algorithms for computing nearest-neighbor decision boundaries. Discrete and Computational Geometry, 2005. 593–604 pp.
39. Davis R.A., Brockwell P.J. Introduction to time series and forecasting 2002. No. 2.

Приложения

Приложение 1. Исходные данные.

Дата	МБ	ГФ	РТС
Янв '00	188,81	1106,23	172,31
Февр '00	191,49	1101,61	170,93
Март '00	254,68	1483,97	231,88
Апр '00	242,85	1450,37	226,87
Май '00	217,05	1208,31	190,21
Июнь '00	194,93	1081,87	171,4
Июль '00	200,22	1213,34	194,09
Авг '00	239,03	1497,59	239,99
Сент '00	204,61	1242,27	199,08
Окт '00	192,22	1182,78	189
Нояб '00	149,84	898,18	143,42
Дек '00	144,39	907,38	143,29
Янв '01	167,74	1107,03	173,53
Февр '01	158,1	1064,08	164,76
Март '01	166,67	1095,96	169,46
Апр '01	178,29	1171,36	180,68
Май '01	207,71	1365,84	208,8
Июнь '01	219,4	1414,65	216,11
Июль '01	200,64	1290,86	196,12
Авг '01	193,72	1356,62	205,41
Сент '01	167,33	1191,67	180,25
Окт '01	187,41	1362,71	204,04
Нояб '01	213,32	1522,84	226,49
Дек '01	237,63	1762,52	260,05
Янв '02	260,77	1984	287,53
Февр '02	262,65	2022,07	290,75
Март '02	302,96	2450,49	350,75
Апр '02	343,33	2708,54	386,1
Май '02	333,61	2754,49	391,26
Июнь '02	304,18	2501,19	353,79
Июль '02	281,81	2306,43	326,23
Авг '02	292,23	2363,64	332,9
Сент '02	289,96	2376,49	334,06
Окт '02	311,41	2559,89	358,65
Нояб '02	321,41	2585,79	361,15
Дек '02	318,91	2566,41	359,07
Янв '03	307,78	2472,79	345,56
Февр '03	341,52	2721,15	383,23
Март '03	325,56	2542,69	360,33
Апр '03	369,77	2953,84	422,37
Май '03	421,08	3221,39	467,1
Июнь '03	455,44	3436,18	503,51

Июль '03	430,3	3109,79	457,02
Авг '03	484	3641,91	530,94
Сент '03	515,17	3900,45	566,62
Окт '03	468,85	3398,23	506,12
Нояб '03	484,73	3539,21	529,27
Дек '03	514,71	3757,15	567,25
Янв '04	551,72	3915,27	611,1
Февр '04	591,09	4296,46	670,14
Март '04	644,64	4821,17	752,66
Апр '04	561,78	4099,08	631,11
Май '04	535,4	3787,34	581,07
Июнь '04	534,84	3807,57	583,32
Июль '04	502,81	3534,06	540,27
Авг '04	549,28	3844,82	584,65
Сент '04	611,03	4149,98	631,65
Окт '04	632,97	4294,87	663,67
Нояб '04	570,76	3987,43	627,98
Дек '04	552,22	3831,97	614,11
Янв '05	575,74	4024,22	637,21
Февр '05	635,38	4474,41	716,42
Март '05	598,04	4186,48	669,07
Апр '05	593,88	4189,79	670,36
Май '05	603,89	4260,47	674,44
Июнь '05	639,98	4554,4	706,38
Июль '05	700,65	5023,84	778,93
Авг '05	784,28	5661,69	882,03
Сент '05	892,5	6458,3	1007,76
Окт '05	842,52	5976,28	934,99
Нояб '05	944,55	6701,53	1037,26
Дек '05	1011	7287,17	1125,6
Янв '06	1171,44	8321,5	1315,96
Февр '06	1320,83	9191,38	1453,44
Март '06	1299,19	8958,64	1434,99
Апр '06	1486,85	10197,14	1657,28
Май '06	1281,5	8866,55	1461,22
Июнь '06	1331,39	9101,18	1494,63
Июль '06	1380,24	9372,74	1551,09
Авг '06	1448,72	9780,59	1626,69
Сент '06	1367,24	9323,57	1549,99
Окт '06	1426,86	9705,26	1613,57
Нояб '06	1550,58	10513,34	1776,68
Дек '06	1693,47	11400,53	1921,92
Янв '07	1656,94	10995,87	1842,93
Февр '07	1655,25	10930,69	1858,14
Март '07	1698,08	11326,33	1935,72

Апр '07	1697,28	11205,07	1935,51
Май '07	1570,34	10370,15	1780,33
Июнь '07	1665,96	11008,32	1897,7
Июль '07	1734,42	11478,56	1993,96
Авг '07	1677,02	11073,54	1919,89
Сент '07	1759,44	11629,43	2071,8
Окт '07	1874,73	12359,45	2223,06
Нояб '07	1850,64	12156,74	2220,11
Дек '07	1888,86	12690,62	2290,51
Янв '08	1574,33	10496,01	1906,97
Февр '08	1660,42	11192,66	2063,94
Март '08	1628,43	10861,12	2053,93
Апр '08	1667,35	11286,48	2122,5
Май '08	1925,24	13090,91	2459,88
Июнь '08	1753,67	12149,79	2303,34
Июль '08	1495,33	10368,79	1966,68
Авг '08	1348,92	9086,68	1646,14
Сент '08	1027,66	6879,83	1211,84
Окт '08	731,96	4616,05	773,37
Нояб '08	611,32	4058,3	658,14
Дек '08	619,53	4174,77	631,89
Янв '09	624,9	4173,08	535,04
Февр '09	666,05	4374,56	544,58
Март '09	772,93	5274,72	689,63
Апр '09	920,35	6264,7	832,87
Май '09	1123,38	7497,44	1087,59
Июнь '09	971,55	6718,52	987,02
Июль '09	1053,3	7213,29	1017,47
Авг '09	1091,98	7544,74	1066,53
Сент '09	1197,2	8350,36	1254,52
Окт '09	1237,18	8648,04	1348,54
Нояб '09	1284,95	8876,82	1377,65
Дек '09	1370,01	9568,83	1444,61
Янв '10	1419,42	9827,37	1473,81
Февр '10	1332,64	9154,18	1410,85
Март '10	1450,15	9955,03	1572,48
Апр '10	1436,04	9811,96	1572,84
Май '10	1332,62	9150,58	1384,59
Июнь '10	1309,31	8979,23	1339,35
Июль '10	1397,12	9648,34	1479,73
Авг '10	1368,9	9364,41	1421,21
Сент '10	1440,3	9776,44	1507,66
Окт '10	1523,39	10463,83	1587,14
Нояб '10	1565,52	10671,83	1597,35
Дек '10	1687,99	11393,51	1770,28

Янв '11	1723,42	11828,74	1870,31
Февр '11	1777,84	12123,66	1969,91
Март '11	1813,59	12411,57	2044,2
Апр '11	1741,84	11893,21	2026,94
Май '11	1666,3	11281,21	1888,6
Июнь '11	1666,59	11544,74	1906,71
Июль '11	1705,18	11544,06	1965,02
Авг '11	1546,05	10506,37	1702,28
Сент '11	1366,54	9303,84	1341,09
Окт '11	1498,6	10191,43	1563,28
Нояб '11	1499,62	10246,37	1540,81
Дек '11	1402,23	9592,15	1381,87
Янв '12	1514,03	10362,82	1577,29
Февр '12	1597,67	10955,11	1734,99
Март '12	1517,34	10362,35	1637,73
Апр '12	1473,5	10086,75	1593,97
Май '12	1306,42	9066,19	1242,43
Июнь '12	1387,52	9559,83	1350,51
Июль '12	1407,02	9609,77	1377,35
Авг '12	1422,91	9718,29	1389,72
Сент '12	1458,26	9986,09	1475,7
Окт '12	1425,7	9769,9	1433,96
Нояб '12	1405,97	9624,04	1436,55
Дек '12	1474,72	10102,06	1526,98
Янв '13	1546,76	10602,53	1622,13
Февр '13	1486,04	10118,12	1534,41
Март '13	1438,57	9766,5	1460,04
Апр '13	1385,88	9392,15	1407,21
Май '13	1350,17	9148,54	1331,43
Июнь '13	1330,46	8993	1275,44
Июль '13	1375,79	9294,64	1313,38
Авг '13	1364,65	9208,07	1290,96
Сент '13	1462,82	9923,83	1422,49
Окт '13	1510,21	10278,04	1480,42
Нояб '13	1479,35	10064,13	1402,93
Дек '13	1504,08	10176,9	1442,73
Янв '14	1454,45	9828,99	1301,02
Февр '14	1444,71	9741	1267,27
Март '14	1369,29	9267,08	1226,1
Апр '14	1306,01	8790,54	1155,7
Май '14	1432,03	9626,25	1295,75
Июнь '14	1476,38	9938,21	1366,08
Июль '14	1379,61	9256,28	1219,36
Авг '14	1400,71	9364,82	1190,23
Сент '14	1411,07	9479,29	1123,72

Окт '14	1488,47	10041,34	1091,44
Нояб '14	1533,68	10308,24	974,27
Дек '14	1396,61	9298,92	790,71
Янв '15	1647,69	10948,57	737,35
Февр '15	1758,97	11756,36	896,63
Март '15	1626,18	10811,84	880,42
Апр '15	1688,34	11299,1	1029,31
Май '15	1609,19	10736,97	968,81
Июнь '15	1654,55	11096,41	939,93
Июль '15	1669	11150,73	858,82
Авг '15	1733,17	11592,75	833,6
Сент '15	1642,97	10906,2	789,73
Окт '15	1711,53	11294,06	845,54
Нояб '15	1771,05	11732,73	847,1
Дек '15	1761,36	11637,91	757,04
Янв '16	1784,92	11874,94	745,3
Февр '16	1840,17	12201,68	768,8
Март '16	1871,15	12372,7	876,2
Апр '16	1953,05	12927,61	951,11
Май '16	1899,01	12451,26	904,33
Июнь '16	1891,09	12360,15	930,77
Июль '16	1944,62	12649,98	927,57
Авг '16	1971,59	12810,85	950,25
Сент '16	1978	12924,39	990,88
Окт '16	1989,64	13012	988,74
Нояб '16	2104,91	13791,86	1029,05
Дек '16	2232,72	14736,86	1152,33
Янв '17	2217,39	14443,65	1164,15
Февр '17	2035,77	13222,07	1099,46
Март '17	1995,9	12953,76	1113,76
Апр '17	2016,71	13074,91	1114,43
Май '17	1900,38	12272,49	1053,3
Июнь '17	1879,5	12146,89	1000,96
Июль '17	1919,53	12345,46	1007,14
Авг '17	2022,22	13016,65	1095,84
Сент '17	2077,19	13417,52	1136,75
Окт '17	2064,31	13300,91	1113,41
Нояб '17	2100,62	13621,36	1131,56
Дек '17	2109,74	13672,33	1154,43
Янв '18	2289,99	15017,86	1282,36
Февр '18	2296,8	15040,05	1285,47
Март '18	2270,98	14911,97	1249,41
Апр '18	2307,02	15270,03	1153,96
Май '18	2302,88	15295,57	1162,98
Июнь '18	2295,95	15277,87	1154,16

Июль '18	2321,11	15427,62	1173,06
Авг '18	2345,85	15618,89	1092,29
Сент '18	2475,36	16597,36	1192,04
Окт '18	2352,71	15723,13	1126,21
Нояб '18	2392,5	15965,66	1126,14
Дек '18	2369,33	15796,11	1068,72
Янв '19	2521,1	16924,6	1214,45
Февр '19	2485,27	16638,24	1188,28
Март '19	2497,1	16729,6	1198,11
Апр '19	2559,32	17266,96	1248,39
Май '19	2665,33	18134,33	1287,09
Июнь '19	2765,85	18763,64	1380,52
Июль '19	2739,5	18672,57	1360,04
Авг '19	2740,04	18580,14	1293,32
Сент '19	2747,18	18609,68	1333,91
Окт '19	2893,98	19743,58	1422,92
Нояб '19	2935,37	20062,09	1438,45
Дек '19	3045,87	20823,93	1548,92
Янв '20	3076,65	20768,01	1517,07
Февр '20	2785,08	18615,21	1299,69
Март '20	2508,81	16549,31	1014,44
Апр '20	2650,56	17472,1	1125,03
Май '20	2734,83	18141,66	1219,76
Июнь '20	2743,2	18212,13	1212,63
Июль '20	2911,57	19189,96	1234,44
Авг '20	2966,2	19598,33	1258,6
Сент '20	2905,81	18958,2	1178,51
Окт '20	2690,59	17409,38	1066,6
Нояб '20	3107,58	20341,87	1281,97
Дек '20	3289,02	21651,06	1387,46
Янв '21	3277,08	21375,21	1367,64
Февр '21	3346,64	21601,18	1411,93
Март '21	3541,72	22832,06	1477,11
Апр '21	3544	22616,62	1485,03
Май '21	3721,63	23838,45	1597,54
Июнь '21	3841,85	24642,75	1653,78
Июль '21	3771,58	24250,51	1625,76
Авг '21	3918,96	25107,6	1684,16
Сент '21	4103,52	26720,38	1777,74
Окт '21	4150	27131,09	1843,83
Нояб '21	3890,59	25453,69	1645,81
Дек '21	3787,26	24747,2	1595,76

Приложение 2. Код Python

```
# -*- coding: utf-8 -*-
pip install pmdarima

#Установка всех необходимых модулей
import pmdarima as pm
import pandas as pd
import csv
import numpy as np
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from sklearn.metrics import mean_squared_error
from statsmodels.tsa.stattools import adfuller
from matplotlib import pyplot as plt
from pmdarima import model_selection
from pmdarima import utils

############################################################################
#Загрузка данных
data = pd.read_csv('/content/sample_data/Индекс-Мосвобиржи.csv', names=['value'], header=0)

# Изменение данных
results = []
with open("/content/sample_data/Индекс-Мосвобиржи.csv") as csvfile:
    reader = csv.reader(csvfile, quoting=csv.QUOTE_NONNUMERIC) # Изменение значений
    for row in reader: # Столбец в список
        results.append(row)
a = np.array(results)
a = a.transpose()
print(a)
don = np.ravel(results)

# Разделим данные на их отдельные части
train, test = model_selection.train_test_split(data, train_size=150)

############################################################################
#Расширенный тест Дики — Фуллера (ADF)
res = adfuller(mydata.value.dropna())
print('Расширенный тест Дики — Фуллера: %f' % res[0])
print('p-значение: %f' % res[1])

plt.rcParams.update({'figure.figsize': (9, 7), 'figure.dpi': 120})

# Исходный временной ряд
fig, axes = plt.subplots(3, 2, sharex=True)
axes[0, 0].plot(data.value);
axes[0, 0].set_title('Исходные данные')
plot_acf(data.value, ax=axes[0, 1])

# Первый порядок дифференцирования
axes[1, 0].plot(data.value.diff());
axes[1, 0].set_title('Первое дифференцирование')
plot_acf(data.value.diff().dropna(), ax=axes[1, 1])

# Второй порядок дифференцирования
axes[2, 0].plot(data.value.diff().diff());
axes[2, 0].set_title('Второе дифференцирование')
```

```
plot_acf(data.value.diff().diff().dropna(), ax=axes[2, 1])

plt.show()

##########################################################################
# Простая модель auto_arima
modl = pm.auto_arima(train, start_p=1, start_q=1, start_P=1, start_Q=1,
                     max_p=5, max_q=5, max_P=5, max_Q=5, seasonal=True,
                     stepwise=True, suppress_warnings=True, D=10, max_D=10,
                     error_action='ignore')

# Создаем прогнозы на будущее
preds, conf_int = modl.predict(n_periods=test.shape[0], return_conf_int=True)

# Оценка прогнозов Среднеквадратической ошибкой
print("Тест RMSE: %.3f" % np.sqrt(mean_squared_error(test, preds)))

##########################################################################
# Наносим на график точки и прогнозы
x_axis = np.arange(train.shape[0] + preds.shape[0])
x_years = x_axis # Исчисление начинается с Января 2000 года и заканчивается в Декабре 2021 Года

plt.plot(x_years[x_axis[:train.shape[0]]], train, alpha=0.75)
plt.plot(x_years[x_axis[train.shape[0]:]], preds, alpha=0.75)  # Прогнозы
plt.scatter(x_years[x_axis[train.shape[0]:]], test,
            alpha=0.4, marker='x')  # Тестовые данные
plt.fill_between(x_years[x_axis[-preds.shape[0]:]],
                 conf_int[:, 0], conf_int[:, 1],
                 alpha=0.1, color='b')
plt.title("Линия прогноза")
plt.xlabel("Месяцы")

##########################################################################
# Сопоставление фактический тестов с прогнозами:
x = np.arange(test.shape[0])
plt.scatter(x, test, marker='x')
plt.plot(x, arima.predict(n_periods=test.shape[0]))
plt.title('Фактические тестовые образцы в сравнении с прогнозами')
plt.show()

# Декомпозиция
figure_kwargs = {'figsize': (6, 6)}  # Установка размера рисунка

# Мультипликативный метод
head_index = 21*12-6
tail_index = 21*12+12
first_index = head_index - tail_index
last_index = head_index
index = np.ravel(data)
timeserie_beer = index[first_index:last_index]

# Декомпозиция набора данных на трендовые, сезонные и случайные части.
decomposed = arima.decompose(index,'multiplicative', m=12)

# Рисунок декомпилированного ряда
axes = utils.decomposed_plot(decomposed, figure_kwargs=figure_kwargs, show=False)
      axes[0].set_title("Сезонная декомпозиция индексов"
```