

EDN: FYSZBG
УДК 339.133

COMPARATIVE ASSESSMENT OF SALES FORECASTING MODELS IN RETAIL

Vladislav R. Rassokhin^{*}, Yulia I. Cherkasova
Siberian Federal University, Krasnoyarsk, Russian Federation

Received 16.05.2024, accepted 03.07.2024

Abstract. This article represents a practical study in the field of data analysis. In the rapidly changing retail market, accurate sales forecasting is crucial for maintaining business competitiveness. The study explores time series forecasting methods and models implemented using machine learning. We analyze sales data from the largest retail network in the Krasnoyarsk region, training the model on a two-year period.

The review of scientific research allowed us to identify the most commonly used methods by analysts, assess the availability of data, and determine a preliminary list of factors influencing sales volume. Our study presents data preprocessing methods, stages of building and using forecast models. Various forecasting methods and machine learning models for building time series are described and compared, such as the Holt-Winters method, the additive linear model Prophet, and the seasonal autoregressive integrated moving average model. To evaluate forecast accuracy, we calculate the mean absolute error, mean squared error, and mean absolute percentage error. This research provides readers with an understanding of sales forecasting possibilities using machine learning models. The obtained results allowed us to justify the choice of the best time series forecast model, enabling more effective inventory management through accurate sales forecasting.

Keywords: sales forecasting, time series models, trade, predictive analytics, forecast evaluation metrics, machine learning (ML).

Citation: Rassokhin, V. R., Cherkasova, Yu. I. (2024). Comparative assessment of sales forecasting models in retail. In: Trade, service, food industry. Vol. 4(3). Pp. 227–240. EDN: FYSZBG



СРАВНИТЕЛЬНАЯ ОЦЕНКА МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ПРОДАЖ В РОЗНИЧНОЙ ТОРГОВЛЕ

Владислав Романович Рассохин^{*}, Юлия Ивановна Черкасова
Сибирский федеральный университет, Красноярск, Российская Федерация

Аннотация. Данная статья представляет собой практическое исследование в области анализа данных. В условиях быстро меняющегося рынка розничной торговли точное прогнозирование продаж имеет решающее значение для сохранения конкурентоспособности бизнеса. В работе рассмотрены методы прогнозирования временных рядов и реализованные на их основе модели с использованием машинного обучения. Авторы анализируют данные о продажах в крупнейшей розничной торговой сети Красноярского края, обучая модель на ежедневных данных двухлетнего временного периода.

Проведенный обзор научных исследований позволил выделить наиболее часто используемые аналитиками методы, оценить наличие данных и определить

предварительный перечень факторов, оказывающих влияние на объем продаж. В статье представлены методы предобработки данных, этапы построения и использования модели прогноза. Описываются и сравниваются различные методы прогнозирования и модели машинного обучения для построения временных рядов, такие как метод Хольта-Винтерса, аддитивная линейная модель Prophet, модель сезонной авторегрессионной интегрированной скользящей средней. Для оценки точности прогноза рассчитывается средняя абсолютная ошибка, среднеквадратическая ошибка и средняя абсолютная процентная ошибка. Настоящее исследование дает читателю представление о возможностях прогнозирования продаж с использованием моделей машинного обучения. Полученные результаты позволили обосновать выбор лучшей модели прогноза временных рядов, что даст возможность более эффективно управлять запасами компании за счет точного прогнозирования объемов продаж.

Ключевые слова: прогнозирование продаж, модели временных рядов, розничная торговля, прогнозная аналитика, метрики оценки прогноза, машинное обучение.

Цитирование: Рассохин, В. Р. Сравнительная оценка моделей прогнозирования продаж в розничной торговле / В. Р. Рассохин, Ю. И. Черкасова // Торговля, сервис, индустрия питания. – 2024. – № 4(3). – С. 227–240. – EDN: FYSZBG



Введение. Точное прогнозирование продаж имеет решающее значение для сохранения конкурентоспособности и прибыльности бизнеса. Постоянные изменения конъюнктуры рынка, складывающиеся под влиянием экономических, политических и социальных факторов, требуют от аналитиков быстрых, адекватных и относительно недорогих методов прогнозирования продажи товаров. Все большее внимание бизнес-сообщества и исследователей привлекают различные методы прогнозирования, использующие современные информационные технологии, в том числе машинное обучение и искусственный интеллект. С развитием электронной коммерции и стремительной трансформацией поведения потребителей, традиционные методы прогнозирования продаж часто не успевают за изменениями и имеют множество ограничений, что вынуждает ритейлеров обращаться к актуальным методам аналитики. Кроме того, доступность больших наборов данных и увеличение мощности вычислительных ресурсов позволяет обучать сложные модели. Однако, на практике есть еще много нерешенных вопросов. Разработка точных моделей прогнозирования продаж остается сложной задачей. По мнению М. Хака [1], аналитикам еще долго предстоит тщательно оценивать сильные и слабые стороны различных методов машинного обучения.

Цель нашего исследования состояла в прогнозировании временных рядов, с использованием различных вариантов машинного обучения и выборе наиболее эффективной модели для прогнозирования продаж товаров предприятия розничной торговли.

Для целей анализа была использована база больших данных одной из самых крупных компаний, занимающихся розничными продажами в Красноярском крае, для тестирования инструментов машинного обучения с целью выявления трендов и прогнозирования физического объема продаж на отдельные товары. Авторы опирались на библиотеки для статистического анализа и визуализации на языке программирования Python.

Одной из наиболее серьезных проблем в прогнозировании продаж является работа со сложными временными рядами данных, на которые, помимо

макроэкономических факторов, может влиять сезонность, недельные и месячные тренды, праздничные дни и т.д. [2]. Сложность применения и использования моделей для прогнозирования временного ряда отмечает К. С Пивкин [3]. Он описывает использование моделей авторегрессии, скользящей средней и модели экспоненциального сглаживания для предсказания количества чеков с последующей оценкой и сравнением прогноза со средним арифметическим значением. В результате оценки модели показывают меньший результат, чем среднее значение тестовых данных. Подобный подход приводится и в статье Е. А. Лясковской и А. М. Коровина [4]. Авторы анализируют не продажи, а спрос, определяют наиболее влиятельные факторы, относящиеся к нему, с помощью построения деревьев решений. В результате исследования сделан вывод, что наибольшее влияние на будущий спрос на рынке дорожно-строительной техники оказали сроки выполнения прошлых заказов (т.е. временной фактор). В работе М. Зожди [5] на основе методов машинного обучения проведено прогнозирование продаж розничного предприятия торговли с использованием клиентских данных (пол и возраст клиента). Автор применял различные модели: k-ближайших соседей (KNN), многослойный перцептрон (MLP) и дерево решений. Р. Вашиштва [6] предлагает прогнозировать продажи с учетом срока нахождения товара в торговом зале или на складе и акцентирует внимание пользы такого прогноза для планирования запасов и выстраивания оптимальной цепочки поставок. Ряд авторов анализируют техническую и операционную стороны вопроса [7–9], обращая внимание на специальное программное обеспечение, огромные объемы данных, необходимость их хранения и защиты, недостаточные навыки аналитиков данных и т.д.

Большинство авторов соглашается с тем, что методы прогнозирования, использующие специализированные библиотеки на языке программирования Python сегодня довольно популярны у исследователей-экономистов, а также широко применяются IT-специалистами, например, в соревнованиях по прогнозированию. В последние годы развитие таких библиотек на языке программирования Python (например, pandas, NumPy, statsmodels и scikit-learn), а также появление новых библиотек (Prophet, PyTorch и TensorFlow [10]) значительно расширили возможности прогнозирования временных рядов. Кроме того, свободный доступ информации в интернете, публикация результатов исследований и кодов программ, рост уровня развития машинного обучения позволяют широко использовать данный инструментарий на практике.

Научная новизна проведенного исследования заключается в обосновании и апробировании методических подходов к разработке прогнозных моделей, позволяющих при их использовании значительно повысить точность прогноза розничных продаж с наименьшими ресурсными и временными затратами. Впервые предложено использовать конкретный метод обработки данных с разбиением переменных через словари и модели для прогноза по каждому предприятию торговли отдельно, что является эффективным методом разбиения данных для множественного прогноза и вносит вклад в практическое знание применений моделей машинного обучения для прогнозирования продаж с использованием большого количества наборов данных. Практическая значимость полученных результатов определяется возможностью их масштабирования для множества номенклатур модели на основе алгоритмов машинного обучения.

Материалы и методы. В нашем исследовании использованы ежедневные данные о продажах розничного предприятия торговли (товар «Бананы») в натуральных объемах реализации – в килограммах. Выбор данного товара обусловлен его круглогодичным наличием в ассортименте, несмотря на сезонность, а также сильной

волатильностью продаж, что, несмотря на усложнение задачи прогноза, усиливает практическую значимость выводов. Временной интервал: период с января 2022 года по апрель 2024 года, плюс тестовый период – апрель 2024 года. Итого в набор данных включено 837 значений с тремя признаками: номер дня недели, номер недели в году, цена продажи. На рис. 1 представлена схема, по которой была проведена разработка прогнозной модели. Данные должны быть выгружены в формате, совместимом со средой разработки (формат CSV, SQL) и приведены с разделителем. Их необходимо загрузить в интегрированную среду разработки (IDE), такую как Jupyter Notebook, VSCode. Процесс предобработки включает удаление выбросов, заполнение пропущенных значений и нормализацию данных [11]. Следующий этап – кросс-валидация – позволяет оценить, насколько хорошо модель будет работать на тренировочных данных, и предотвратить переобучение.

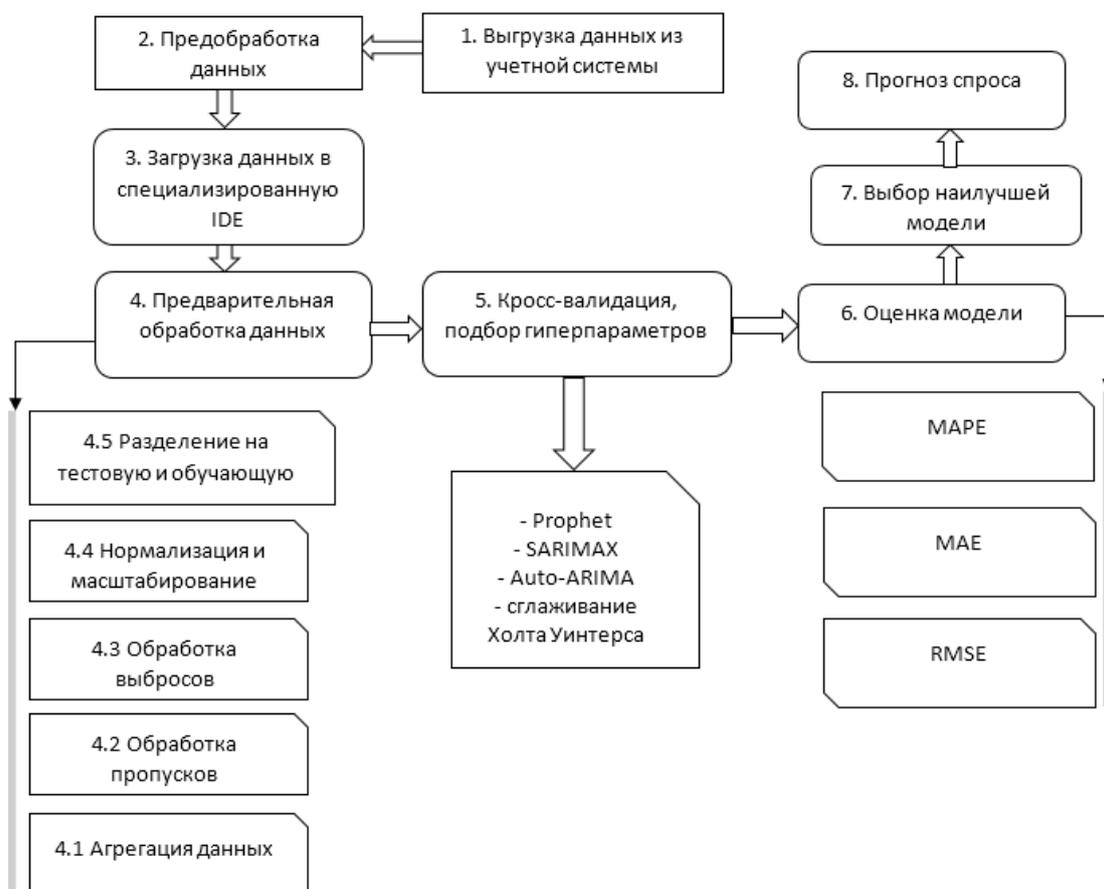


Рисунок 1. Этапы разработки модели прогнозирования [составлено авторами]
 Figure 1. Stages of application of the forecasting model [author's construction]

Для подбора гиперпараметров каждой модели используется инструмент библиотеки Scikit-learn - GridSearchCV [12]. Оценку модели производили на тестовых данных, которые не применялись для расчета и обучения моделей. Авторы использовали для оценки метрики MAPE, MAE и RMSE.

MAPE, или средняя абсолютная процентная ошибка, является одной из метрик оценки точности прогнозов. Ее вычисляют как среднее значение абсолютных процентных ошибок между фактическими и прогнозируемыми значениями. MAPE удобна тем, что выражается в процентах, а это делает ее легкой для интерпретации и

сравнения между различными моделями прогнозирования [13]. Однако MAPE может быть чувствительной к нулевым значениям в данных. Формула расчета

$$MAPE = 100 \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (1)$$

где \hat{y}_i – прогнозное значение;

y_i – фактическое значение;

n – размер выборки.

Среднеквадратическая ошибка (RMSE) – это метрика оценки качества модели или алгоритма, которая измеряет среднюю величину ошибки между предсказанными и прогнозными значениями. Чем меньше значение метрики, тем лучше модель предсказывает значения. RMSE рассчитывают как квадратный корень из среднего значения квадратов ошибок между предсказанными и прогнозными значениями. Формула для расчета RMSE выглядит следующим образом [14]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (2)$$

MAE (средняя абсолютная ошибка) – это метрика, которая измеряет среднюю величину ошибок в предсказаниях модели. Она рассчитывается как среднее значение абсолютных разностей между предсказанными и прогнозными значениями.

Формула для расчета MAE [15]:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|. \quad (3)$$

MAE популярна, потому что она:

- легко интерпретируется, дает представление о среднем размере ошибки;
- менее чувствительна к экстремальным значениям (выбросам), чем RMSE, который может быть искажен несколькими крупными ошибками;
- вычисляется быстрее, чем RMSE, особенно для больших наборов данных.

На следующем этапе авторами были выбраны модели, успешно зарекомендовавшие себя в работах предыдущих исследователей (ARIMA, Prophet, модель Холта-Уинтерса) для тестирования и выбора наилучшей. Остановимся на характеристиках моделей и обосновании их выбора.

ARIMA – это модель авторегрессии и интегрированной скользящей средней. Эти модели достаточно гибкие и описывают совокупность характеристик временных рядов. Авторегрессия означает, что каждое значение ряда находится в линейной зависимости от предыдущих значений; скользящая средняя предполагает, что в ошибках модели в предшествующие периоды сосредоточена информация обо всей предыстории ряда [16].

SARIMA (Сезонная ARIMA) – это расширение модели ARIMA, которое включает сезонные компоненты. Модель SARIMA описывается параметрами

$$(p,d,q) \times (P,D,Q,S) (p,d,q) \times (P,D,Q,S),$$

где p,d,q,p,d,q – параметры несезонной части;

P,D,Q,P,D,Q – параметры сезонной части модели;

SS – длина сезонного периода.

SARIMAX (Сезонная ARIMA с экзогенными переменными) – последняя модификация данной модели, используемая нами в расчетах. Она расширяет SARIMA путем добавления возможности учитывать экзогенные переменные «X», то есть внешние переменные или регрессоры, которые могут влиять на зависимую переменную временного ряда. Это позволяет модели учитывать дополнительную информацию, которая может улучшить точность прогноза.

Метод Auto-ARIMA является автоматизированным методом выбора оптимальных параметров модели ARIMA на основе исторических данных временного ряда. Этот метод основан на принципе минимизации информационного критерия Акаике (AIC) [17] и может быть полезным в случае, когда нет явных представлений о структуре временного ряда. Однако, его результативность может быть ограничена в случае наличия сильной сезонности или вариабельности данных.

Вторая рассматриваемая в данном исследовании модель – Prophet [18, 19]. Это аддитивная модель, которая состоит из следующих компонентов:

1. Тренд используется для моделирования общего направления изменения временного ряда. Тренд может быть линейным или нелинейным.

2. Сезонность используется для моделирования периодических изменений в временном ряде, таких как ежегодные, еженедельные или ежедневные сезонные колебания.

3. Праздники используются для моделирования эффектов праздников и других нерегулярных событий на временной ряд.

Модель в Prophet может быть представлена как

$$y(t) = trend(t) + seasonality(t) + holidays(t) + error(t), \quad (4)$$

где $y(t)$ – значение временного ряда в момент времени t ,

$trend(t)$ – трендовая компонента;

$seasonality(t)$ – сезонная компонента;

$holidays(t)$ – компонента праздников;

$error(t)$ – остаточное значение.

Модель в Prophet также использует следующие техники:

- Piecewise linear trend: для моделирования нелинейного тренда;
- Fourier series: для моделирования сезонности;
- Regression: для моделирования эффектов праздников.

В целом модель в Prophet позволяет эффективно прогнозировать временные ряды с нелинейными трендами и сезонными колебаниями. Нами эта модель используется по примеру Дж. Тейлора и Б. Летамы [20], которые в своей работе предложили аддитивную модель для прогнозирования временных рядов и получили хорошие результаты.

Модель Холта-Уинтерса, также известная как экспоненциальное сглаживание тройного уровня, является одним из методов прогнозирования временных рядов, который учитывает не только тренд (как в модели Холта), но и сезонность [21]. Эта модель особенно полезна для рядов с явной сезонной компонентой и может быть, как аддитивной, так и мультипликативной в зависимости от характера сезонности.

Модель Холта-Уинтерса включает в себя три компоненты (тренд, сезонность, уровень), которые обновляются на каждом временном интервале (точке данных), что позволяет модели адаптироваться к изменениям в данных.

1. Тренд: оценивается тренд, который показывает, увеличивается ли или уменьшается значение временного ряда в долгосрочной перспективе.

2. Сезонность: оценку сезонной компоненты, которая отражает периодические колебания, связанные с сезонными факторами.

3. Уровень: на каждом шаге модель оценивает текущий базовый уровень данных, который отражает среднее значение временного ряда в данный момент.

На каждом временном интервале в модели используется последние доступные данные для обновления оценок. Это делается с помощью формул экспоненциального сглаживания, которые включают параметры сглаживания для уровня, тренда и сезонности. Данные параметры контролируют, насколько сильно новые данные влияют на оценки модели.

Полученные результаты. В этом разделе представлен прогноз физического объема продаж по товару «Бананы». Необходимые данные были собраны из учетных систем одного из предприятий торговли региональной торговой сети, охватывают двухлетнюю историю продаж данного товара. Динамика продаж за 837 дней представлена на рис. 2 и 3.

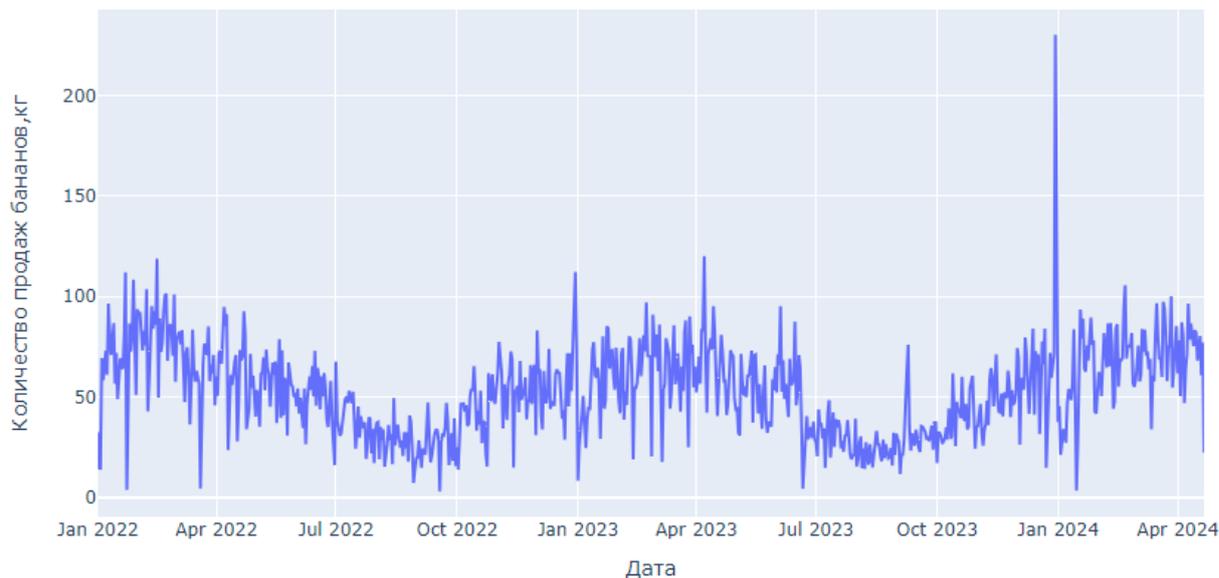


Рисунок 2. Динамика продажи товаров (бананы) за анализируемый период, кг
Figure 2. Sales dynamics of products (bananas) for the analyzed period, kg

На рис. 2 заметна сезонность продаж. Хотя в отдельные дни наблюдается значительная волатильность (о чем было сказано ранее), в среднем объемы продаж снижаются во второй половине лета – начале осени.

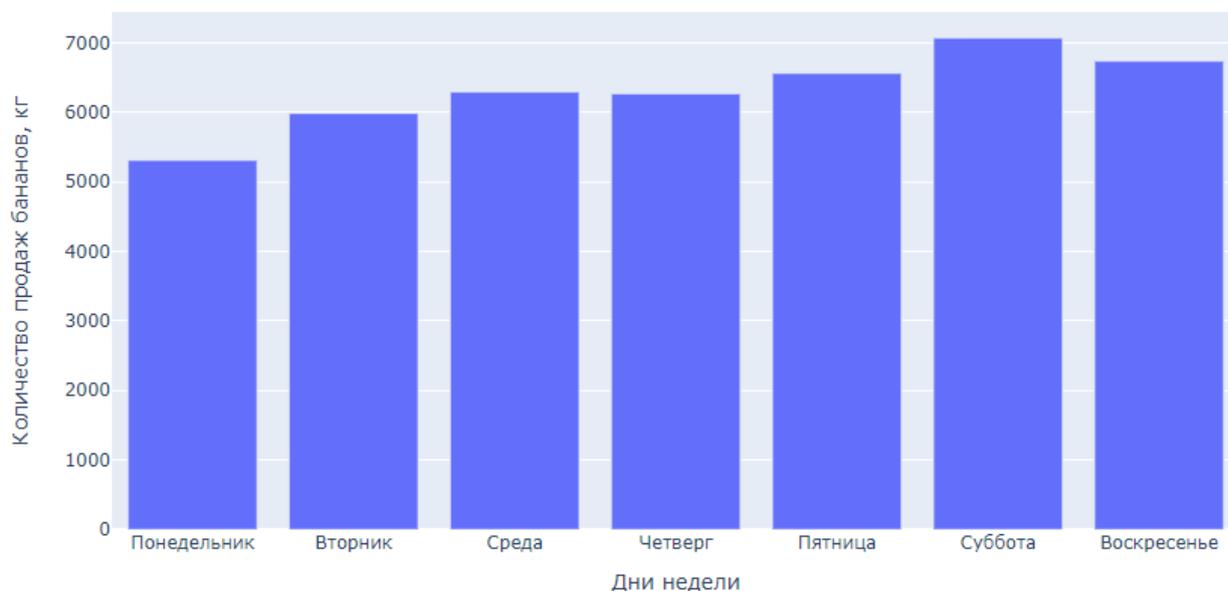


Рисунок 3. Распределение продажи товаров по дням недели за анализируемый период
Figure 3. Distribution of product sales by days of the week for the analyzed period

Представленная на рис. 3 группировка данных по суммированным объемам продаж по дням недели демонстрирует недельную сезонность с минимальными продажами в понедельник и максимальными в субботу.

Далее, для лучшего понимания необходимости использования конкретной модели для наших целей, представлена описательная характеристика данных временного ряда (табл. 1).

Таблица 1. Описательные характеристики данных временного ряда модели
Table 1. Descriptive characteristics of time series data

Показатель	Значение
Минимальное значение продаж, кг	20,7
Максимальное значение продаж, кг	230,4
Математическое ожидание	52,8
Дисперсия	505,1
Стандартное отклонение	22,4
Экссесс	3,8

Дисперсия и стандартное отклонение показывают, что данные имеют значительный разброс относительно среднего значения. Экссесс, равный 3,8, указывает на то, что распределение имеет более острый пик и более толстые хвосты по сравнению с нормальным распределением. Исходя из этих характеристик, можно предположить, что данные имеют положительную асимметрию (правостороннее смещение) и более выраженный пик, чем у нормального распределения. Это может указывать на наличие выбросов или на то, что данные не следуют нормальному распределению, а, возможно, подчиняются гамма-распределению, данная информация поможет при подборе гиперпараметров.

SARIMAX позволяет моделировать сложные временные ряды с учетом как обычной динамики, так и сезонности, а также с учетом внешних регрессоров. Результаты прогнозирования продаж на 28 дней на основе модели SARIMAX представлены на рис. 4.

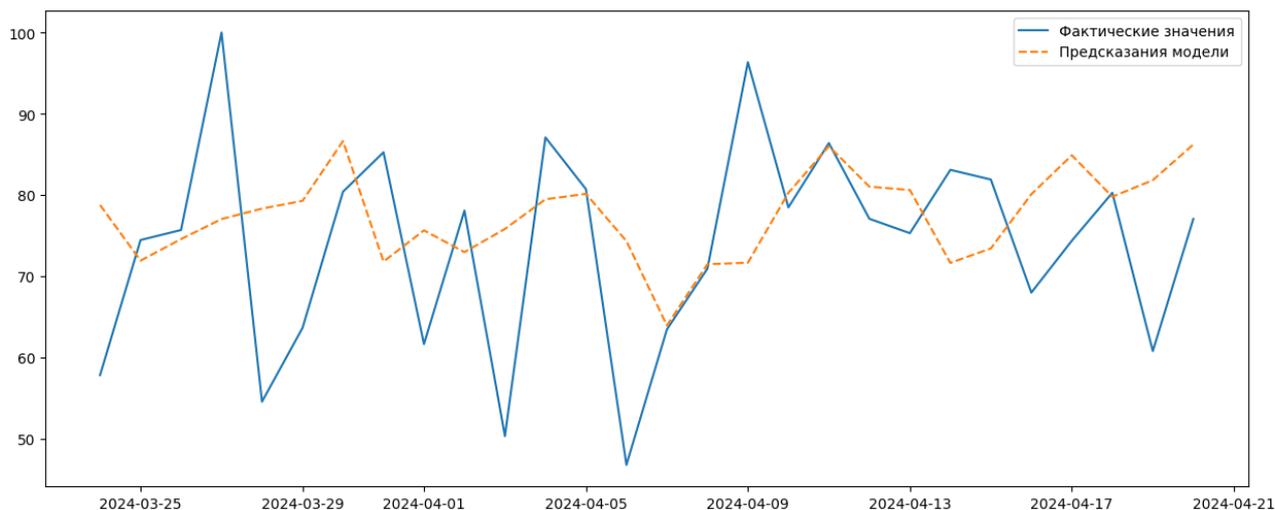


Рисунок 4. Результаты прогноза продаж бананов по модели SARIMAX, кг
Figure 4. Sales forecast results for bananas using the SARIMAX model, kg

Рассчитанные значения метрик по указанной модели представлены в табл. 2.

Таблица 2. Сравнительная оценка моделей прогнозирования объемов продаж
Table 2. Comparative assessment of sales volume forecasting models

Метрика	Среднее арифметическое значение продаж	SARIMAX (с учетом регрессоров)	Auto-ARIMA	Экспоненциальное сглаживание Холта-Уинтерса	Prophet (с учетом регрессоров)
RMSE	13,4	13.8	12.7	21,3	13,9
MAE	9,9	10.6	9.9	17,2	11,35
MAPE, %	15,6	16	14.9	24	15,9

Использованное интегрирование порядка 1 указывает на однократное дифференцирование временного ряда для обеспечения стационарности. Параметр сезонной составляющей определен как 12 (годовой). Для построения модели использовались регрессоры: цена, номер дня недели, номер недели в году. Коэффициент признака «Цена» составил $-0,1693$, что означает ожидаемую отрицательную связь между ценой и продажами. Коэффициент «Номер дня недели» имеет значение, равное $1,9501$, и указывает на положительную связь между конкретным днем недели и продажами.

Далее были использованы еще 3 модели прогнозирования: Auto-ARIMA, модель Холта-Уинтерса и Prophet. Auto-ARIMA с возможностью автоматического подбора гиперпараметров стремится минимизировать критерий Акаике, значение которого может указывать на то, что если к расчету модели прибавить свободный коэффициент, то ее прогнозные значения улучшатся. Модель экспоненциальное сглаживание Холта-Уинтерса использует экспоненциальное сглаживание для обновления и прогнозирования трех компонентов (тренда, уровня и сезонности). Она учитывает веса для последних значений временного ряда, придавая больший вес более поздним данным. Однако, в соответствии со значениями наших метрик, в данном исследовании эта модель показала наихудший результат (табл. 2).

Прогноз временного ряда с помощью модели Prophet представлен на рис. 5. На графике отображается прогнозируемое значение (синяя линия) вместе с доверительным интервалом (заштрихованная область), который отражает уровень неопределенности в прогнозе. Это позволяет оценить как надежность прогноза, так и риски, связанные с предсказанными значениями временного ряда.



Рисунок 5. Прогноз модели Prophet на 28 дней с доверительным интервалом

Figure 5. Prophet model forecast for 28 days with a confidence interval

Для сопоставления результатов моделей и выявления наиболее точного прогноза сравним метрики, обозначенные нами в предыдущей части статьи (табл. 2).

Проведенные расчеты позволяют сделать вывод, что модель метода Auto-ARIMA на основе минимизации информационного критерия Акаике демонстрирует наилучшую точность прогнозирования по всем рассматриваемым метрикам. Модели SARIMAX и Prophet (с учетом регрессоров) также показывают приемлемую точность прогнозирования, близкую к лучшим результатам, но в то же время они уступают средней арифметической.

Обсуждение. При интерпретации полученных результатов стоит учитывать, что временной ряд имеет довольно высокую дисперсию. Такое значение дисперсии может указывать на высокий уровень неопределенности или неясности в прогнозируемых значениях. Возможно, требуется более сложная или гибкая модель, чтобы учесть изменчивость и разнообразие данных. Предоставленные модели могут применяться в практической деятельности организации торговли для реального прогноза продаж сезонных товаров с последующим планированием товарных запасов, страхового запаса, а также для разработки графиков доставки товаров от поставщиков. Отметим, что данные методы и модели можно применить как для прогнозирования выручки по группе товаров, так и в целом по предприятию торговли для расчета планируемой выручки. Подобные исследования проводились О. А. Милюшенко [22], где автор использовал линейную регрессию для прогнозирования выручки предприятия, и И. А. Вахрушевым [23], который строил различные модели прогноза с их последующей оценкой. По нашему мнению, данный опыт может быть полезен будущим исследователям при решении похожих прикладных задач.

Выводы и дискуссионные вопросы. Результаты проведенного исследования позволяют сделать вывод о возможности использовании всех рассмотренных моделей машинного обучения для прогнозирования продаж. В нашем случае, согласно выбранным метрикам, наилучшие результаты показала модель Auto-ARIMA. Однако отметим, что нельзя однозначно рекомендовать данную модель как единственную. Представленный в статье порядок (этапы) разработки модели позволит, с учетом предложенных нами критериев, сравнить и выбрать оптимальную модель машинного обучения в зависимости от конкретной цели прогнозирования.

Прогнозирование продаж – один из важных аналитических функционалов для аналитиков данных, работающих в розничном ритейле. По сути в современном мире для онлайн продаж это стало одним из ключевых этапов в планировании цепочек поставок [24]. Однако, к сожалению, перечисленные нами инструменты используются нечасто. На наш взгляд, основная причина – отсутствие понимания необходимости применения новых цифровых методов аналитики и прогнозирования. Также отметим отсутствие полноценной базы или недостаток систематизированных данных о клиентах, их поведении, отсутствие возможности сбора дополнительных параметров, кроме как за счет карт лояльности и развития в сторону интернет-бизнеса. По мнению авторов, наука о данных с каждым годом набирает популярность, а методы прогнозирования становятся более гибкими и эффективными. Появляются новые публикации с результатами апробаций разных инструментов моделирования. Например, исследователи Раджат Сен и Ичен Чжоу [25] предлагают принципиально новый подход к прогнозированию временных рядов с использованием модели, основанной исключительно на декодере. Они описывают архитектуру модели, ее основные компоненты и принцип работы. Эти тенденции раскрывают новые перспективы для проведения дальнейших исследований в прогнозировании продаж для предприятий торговли.

Что касается нашего исследования, авторы видят дальнейшее его развитие в расширении товаров до нескольких категорий, а затем, до полного ассортимента, добавление новых зависимых факторов (в том числе с расчетом через модели), автоматизированное сравнение нескольких моделей для одного товара и выбор лучшей.

Библиографический список

1. Haque, M. S. Retail demand forecasting: a comparative study for multivariate time series [Electronic resource] / M. S. Haque, M. Sh. Amin, J. Miah. – URL: <https://arxiv.org/abs/2308.11939> (Date of access: 12.04.2024). DOI: 10.48550/arXiv.2308.11939.
2. Пивкин, К. С. Корреляционный анализ факторов влияния на покупательский спрос розничного магазина как этап формирования модели прогнозирования и управления запасами [Текст] / К. С. Пивкин // Вестник Удмуртского университета. Серия Экономика и право. – 2016. – Т. 26, вып. 3. – С. 40–50.
3. Пивкин, К. С. Прогнозирование ключевых показателей розничной сети во времени [Текст] / К. С. Пивкин // Вестник Пермского университета. Серия: Экономика. – 2017. – Т. 12, № 4. – С. 592–606. DOI: 10.17072/1994-9960-2017-4-592-608.
4. Коровин, А. М. Прогнозирование спроса на рынке дорожно-строительной техники с использованием инструментов интеллектуального анализа данных [Текст] / А. М. Коровин // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2022. – Т. 22, № 3. – С. 117–131.
5. Demand forecasting based machine learning algorithms on customer information: an applied approach [Electronic resource] / M. Zohdi, M. Rafiee, V. Kayvanfar, A. Salamiraad // International Journal of Information Technology. – 2022. – No 14(3). – URL: https://www.researchgate.net/publication/358567204_Demand_forecasting_based_machine_learning_algorithms_on_customer_information_an_applied_approach (Date of access: 12.05.2024). DOI: 10.1007/s41870-022-00875-3.
6. Product age based demand forecast model for fashion retail [Electronic resource] / R. K. Vashishtha, V. Burman, R. Kumar [et al.]. – URL: <https://arxiv.org/abs/2007.05278> (Date of access: 12.05.2024). DOI: 10.48550/arXiv.2007.05278.
7. Machine-learning algorithm for demand forecasting problem [Electronic resource] / D. T. Wiyanti, I. Kharisudin1, A. B. Setiawan, A. K. Nugroho // Journal of Physics: Conference Series. – URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1918/4/042012/pdf> (Date of access: 12.05.2024). DOI: 10.1088/1742-6596/1918/4/042012.
8. Time series analysis: forecasting and control [Electronic resource] / G. E. P. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung. – URL: http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf (Date of access: 12.05.2024).
9. Ferreira, K. J. Analytics for an online retailer: demand forecasting and price optimization [Electronic resource] / K. J. Ferreira, B. H. A. Lee, D. Simchi-Levi // Manufacturing & Service Operations Management. – URL: <http://hdl.handle.net/1721.1/101783> (Date of access: 12.05.2024). DOI: 10.1287/msom.2015.0561.
10. Teng, S. Route planning method for cross-border e-commerce logistics of agricultural products based on recurrent neural network [Electronic resource] / S. Teng // Soft Computing. – URL: https://www.researchgate.net/publication/351508064_Route_planning_method_for_cross-border_e-commerce_logistics_of_agricultural_products_based_on_recurrent_neural_network (Date of access: 10.05.2024). DOI:10.1007/s00500-021-05861-8.

11. Старовойтов, В. В. Нормализация данных в машинном обучении [Текст] / В. В. Старовойтов, Ю. И. Голуб // Информатика. – 2021. – № 18(3). – С. 83–96. DOI: 10.37661/1816-0301-2021-18-3-83-96.
12. Анафиев, А. С. Обзор подходов к решению задачи оптимизации гиперпараметров для алгоритмов машинного обучения [Текст] / А. С. Анафиев, А. С. Карюк // Таврический вестник информатики и математики. – 2022. – № 2(55). – С. 30–37.
13. Tofallis, C. A better measure of relative prediction accuracy for model selection and model estimation [Text] / C. Tofallis // Journal of the Operational Research Society. – 2015. – No 66(8). – Pp. 1352–1362. DOI: 10.1057/jors.2014.103.
14. Hodson, T. O. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not [Electronic resource] / T. O. Hodson // Geoscientific Model Development. – 2022. – Vol. 15, Is. 14. – Pp. 5481–5487. – URL: <https://gmd.copernicus.org/articles/15/5481/2022/> (Date of access: 10.05.2024). DOI: 10.5194/gmd-15-5481-2022.
15. Bickel, P. Mathematical Statistics: Basic Ideas and Selected Topics [Electronic resource] / P. Bickel, K. Doksum // Journal of the American Statistical Association. – 2007. – No 56(4). – URL: https://www.researchgate.net/publication/236736826_Mathematical_Statistics_Basic_Ideas_and_Selected_Topics (Date of access: 10.05.2024). DOI: 10.2307/2286373.
16. Миролюбова, А. А. ARIMA – прогнозирование спроса производственного предприятия [Текст] / А. А. Миролюбова, А. Д. Ермолаев // Современные наукоемкие технологии. Региональное приложение. – 2021. – № 2(66). – С. 50–55.
17. Kolassa, S. Combining exponential smoothing forecasts using Akaike weights [Electronic resource] / S. Kolassa // International Journal of Forecasting. – 2011. – Vol. 27, Is. 2. – Pp. 238–251. – URL: <https://www.sciencedirect.com/science/article/abs/pii/S0169207010001032> (Date of access: 13.05.2024).
18. VanBerlo, B. Univariate Long-Term Municipal Water Demand Forecasting [Electronic resource] / B. VanBerlo, M. A. S. Ross, D. Hsia. – URL: <https://arxiv.org/abs/2105.08486> (Дата обращения: 13.05.2024). DOI: 10.48550/arXiv.2105.08486.
19. Hoffmann, D. Impact of HPO on AutoML Forecasting Ensembles [Electronic resource] / D. Hoffmann. – URL: <https://arxiv.org/abs/2311.04034> (Date of access: 13.05.2024). DOI: 10.48550/arXiv.2311.04034.
20. Taylor, S. J. Forecasting at scale [Electronic resource] / S. J. Taylor, B. Letham. – URL: https://www.researchgate.net/publication/344989540_Forecasting_at_scale (Date of access: 13.05.2024). DOI: 10.7287/peerj.preprints.3190v2.
21. Выбор метода прогнозирования основных статистических показателей работы ГБУЗ «НИИ СП им. Н. В. Склифосовского Департамента здравоохранения города Москвы» [Текст] / Б. Л. Курилин, В. Я. Киселевская-Бабинина, Н. А. Карасёв [и др.] // Журнал им. Н. В. Склифосовского «Неотложная медицинская помощь». – 2019. – № 8(3). – С. 246–256.
22. Милюшенко, О. А. Прогнозирование выручки от продаж торгового предприятия [Текст] / О. А. Милюшенко, А. Г. Пупкова, А. И. Ковалев // Вестник Сибирского института бизнеса и информационных технологий. – 2020. – № 4(36). – С. 58–64.
23. Вахрушев, И. А. Прогнозирование годовой выручки российских компаний крупного и среднего бизнеса отрасли торговля [Текст] / И. А. Вахрушев // Научный журнал НИУ ИТМО. Серия Экономика и экологический менеджмент. – 2020. – № 3. – С. 45–51.

24. Murray, P. W. Forecasting Supply Chain Demand by Clustering Customers [Electronic resource] / P. W. Murray, B. Agard, M. A. Barajas // IFAC-PapersOnLine. – 2015. – Vol. 48, Is. 3. – Pp. 1834–1839. – URL: <https://www.sciencedirect.com/science/article/pii/S2405896315005923> (Date of access: 05.10.2024).

25. A decoder-only foundation model for time-series forecasting [Electronic resource] / A. Das, W. Kong, R. Sen, Y. Zhou. – URL: <https://arxiv.org/abs/2310.10688> (Date of access: 05.10.2024). DOI: 10.48550/arXiv.2310.10688.

References

1. Haque, M. S., Amin, M. Sh., Miah J. (2023). Retail demand forecasting: a comparative study for multivariate time series. [Electronic source] URL: <https://arxiv.org/abs/2308.11939> (Date of access: 12.04.2024). DOI: 10.48550/arXiv.2308.11939.

2. Pivkin, K. S. (2016). Correlation analysis of factors influencing consumer demand of a retail store as a stage in the formation of a forecasting and inventory management model. *Argancha University Bulletin. Economics and Law Series*, Vol. 26, Is. 3, 40–50.

3. Pivkin, K. S. (2017). Forecasting key indicators of a retail network over time. *Bulletin of the Perm University. Series: Economics*, Vol. 12, Is. 4, 592–606. DOI: 10.17072/1994-9960-2017-4-592-608.

4. Korovin, A. M. (2022). Forecasting demand in the road construction equipment market using data mining tools. *Bulletin of the South Ural State University. Series: Computer technology, control, radio electronics*, Vol. 22, Is. 3, 117–131.

5. Zohdi, M., Rafiee, M., Kayvanfar, V., Salamiraad, A. (2022). Demand forecasting based machine learning algorithms on customer information: an applied approach. *International Journal of Information Technology*, 14(3). [Electronic source] URL: https://www.researchgate.net/publication/358567204_Demand_forecasting_based_machine_learning_algorithms_on_customer_information_an_applied_approach (Date of access: 12.05.2024). DOI: 10.1007/s41870-022-00875-3.

6. Vashishtha, R. K., Burman, V., Kumar, R., Sethuraman, S., Sekar, A. R., Ramanan, S. (2020). Product age based demand forecast model for fashion retail. [Electronic source] URL: <https://arxiv.org/abs/2007.05278> (Date of access: 10.05.2024). DOI: 10.48550/arXiv.2007.05278.

7. Wiyanti, D. T., Kharisudin1, I., Setiawan, A. B., Nugroho, A. K. (2021). Machine-learning algorithm for demand forecasting problem. *Journal of Physics: Conference Series*. [Electronic source] URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1918/4/042012/pdf> (Date of access: 11.05.2024). DOI: 10.1088/1742-6596/1918/4/042012.

8. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., Ljung, G. M. (1976). Time series analysis: forecasting and control. [Electronic source] URL: http://repo.darmajaya.ac.id/4781/1/Time%20Series%20Analysis_%20Forecasting%20and%20Control%20%28%20PDFDrive%20%29.pdf (Date of access: 12.05.2024).

9. Ferreira, K. J., Lee, B. H. A., Simchi-Levi, D. (2015). Analytics for an online retailer: demand forecasting and price optimization. *Manufacturing & Service Operations Management*. [Electronic source] URL: <http://hdl.handle.net/1721.1/101783> (Date of access: 12.05.2024). DOI: 10.1287/msom.2015.0561.

10. Teng, S. (2021). Route planning method for cross-border e-commerce logistics of agricultural products based on recurrent neural network. *Soft Computing*. [Electronic source] URL: https://www.researchgate.net/publication/351508064_Route_planning_method_for_cross-border_e-commerce_logistics_of_agricultural_products_based_on_recurrent_neural_network (Date of access: 10.05.2024). DOI:10.1007/s00500-021-05861-8.

11. Starovoitov, V. V, Golub, Yu. I. (2021). Data normalization in machine learning. *Computer science*, 18(3), 83–96. DOI: 10.37661/1816-0301-2021-18-3-83-96.
12. Anafiev, A. S., Karyuk, A. S. (2022). Review of approaches to solving the problem of optimization of hyperparameters for machine learning algorithms. *Tauride Bulletin of Computer Science and Mathematics*, 2(55). 30–37.
13. Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society*, 66(8), 1352–1362. DOI: 10.1057/jors.2014.103.
14. Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not *Geoscientific Model Development*, Vol. 15, Is. 14, 5481–5487. [Electronic source] URL: <https://gmd.copernicus.org/articles/15/5481/2022/> (Date of access: 10.05.2024). DOI: 10.5194/gmd-15-5481-2022.
15. Bickel, P., Doksum, K. (2007). *Mathematical Statistics: Basic Ideas and Selected Topics*. *Journal of the American Statistical Association*, 56(4). [Electronic source] URL: https://www.researchgate.net/publication/236736826_Mathematical_Statistics_Basic_Ideas_and_Selected_Topics (Date of access: 10.05.2024). DOI: 10.2307/2286373.
16. Mirolyubova, A. A., Ermolaev, A. D. (2021). ARIMA – forecasting the creativity of a manufacturing enterprise. *Modern high technology. Regional application*, 2(66), 50–55.
17. Kolassa, S. (2011). Combining exponential smoothing forecasts using Akaike weights. *International Journal of Forecasting*, Vol. 27, Is. 2, 238–251. [Electronic resource] URL: <https://www.sciencedirect.com/science/article/abs/pii/S0169207010001032> (Date of access: 13.05.2024).
18. VanBerlo, B, Ross, M. A. S., Hsia, D. (2021). Univariate Long-Term Municipal Water Demand Forecasting. [Electronic source] URL: <https://arxiv.org/abs/2105.08486> (Date of access 13.05.2024). DOI: 10.48550/arXiv.2105.08486.
19. Hoffmann, D. (2023). Impact of HPO on AutoML Forecasting Ensembles. [Electronic source] URL: <https://arxiv.org/abs/2311.04034> (Date of access: 13.05.2024). DOI: 10.48550/arXiv.2311.04034.
20. Taylor, S. J., Letham, B. (2018). Forecasting at scale. [Electronic source] URL: https://www.researchgate.net/publication/344989540_Forecasting_at_scale (Date of access: 13.05.2024). DOI: 10.7287/peerj.preprints.3190v2.
21. Kurilin, B. L., Kiselevskaya-Babinina, V. Ya., Karasev, N. A., Kiselevskaya-Babinina, I. V., Kislukhina, E. V., Vasiliev, V. A. (2019). Selection of a method for predicting the main statistical indicators of the work of the State Budgetary Institution “Research Institute of SP named after N. V. Sklifosovsky Department of Health of the City of Moscow”. *Journal named after N. V. Sklifosovsky “Emergency medical care”*, 8(3), 246–256.
22. Milushenko, O. A., Pupkova, A. G., Kovalev, A. I. (2020). Forecasting revenue from sales of a trading enterprise. *Bulletin of the Siberian Institute of Business and Information Technology*, 4(36), 58–64.
23. Vakhrushev, I. A. (2020). Forecasting the annual revenue of Russian companies of large and medium-sized businesses in the trade industry. *Scientific journal of the ITMO Research Institute. Economics and Environmental Management Series*, 3, 45–51.
24. Murray, P. W., Agard, B., Barajas, M. A. (2015). Forecasting Supply Chain Demand by Clustering Customers. *IFAC-PapersOnLine*, 48(3), 1834–1839. [Electronic resource] URL: <https://www.sciencedirect.com/science/article/pii/S2405896315005923> (Date of access: 05.10.2024).
25. Das, A., Kong, W., Sen, R., Zhou, Y. (2023). A decoder-only foundation model for time-series forecasting. [Electronic resource] URL: <https://arxiv.org/abs/2310.10688> (Date of access: 05.10.2024). DOI: 10.48550/arXiv.2310.10688.