

УДК 330.4 JEL C02

UDC 330.4 JEL C02

5.2.2. Математические, статистические и инструментальные методы в экономике

5.2.2. Mathematical, statistical and instrumental methods in economics

ГИБРИДНАЯ МОДЕЛЬ ОПЕРАТИВНОГО ПРОГНОЗА АССОРТИМЕНТНОЙ ПОЗИЦИИ ПОРТФЕЛЯ МОНОПРОДУКТОВОЙ ТОРГОВОЙ КОМПАНИИ

HYBRID MODEL OF OPERATIONAL FORECAST OF THE ASSORTMENT POSITION OF THE PORTFOLIO OF A SINGLE-PRODUCT TRADING COMPANY

Попова Маргарита Игоревна
ассистент
ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина»,
Краснодар, Россия

Popova Margarita Igorevna
assistant
Kuban State Agrarian University named after I. T. Trubilin, Krasnodar, Russia

Статья посвящена микроуровневому анализу, акцентированному на прогнозировании ежедневного спроса позиции ассортиментного портфеля, что позволяет монопродуктовой торговой компании оперативно адаптироваться к динамике рыночной конъюнктуры, оптимизировать уровень товарных запасов и минимизировать операционные издержки. Прогнозная информация получена на основе авторской разработки «Информационная система прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS», которая реализует алгоритм CATSF, представляющий собой новый подход, основанный на интеграции методов машинного обучения, для реализации автоматизированного подхода к обучению линейной клеточно-автоматной модели. Данный подход позволяет избежать субъективных оценок экспертов и снизить ошибку прогнозирования благодаря применению машинного обучения для определения оптимального количества термов в процессе обучения ЛКА. Это способствует повышению объективности и точности моделирования, особенно в части, где экспертная оценка может быть затруднена или предвзята. Гибридная прогнозная модель апробирована на реальных данных для получения оперативного прогноза ассортиментной позиции портфеля монопродуктовой торговой компании

The article is devoted to a micro-level analysis focused on forecasting the daily demand for product portfolio items, which allows a single-product trading company to quickly adapt to the dynamics of market conditions, optimize inventory levels and minimize operating costs. The predictive information was obtained on the basis of the author's development "Information system for predicting time series dynamics based on a cellular automaton with learning automation CAFTS", which implements the CATSF algorithm, which is a new approach based on the integration of machine learning methods to implement an automated approach to learning a linear cellular automaton model. This approach avoids subjective expert assessments and reduces prediction error by using machine learning to determine the optimal number of terms in the learning process of a linear cellular automaton. This helps to increase the objectivity and accuracy of modeling, especially in the part where expert assessment may be difficult or biased. The hybrid predictive model has been tested on real data to obtain an operational forecast of the product range of the portfolio of a single-product trading company

Ключевые слова: АССОРТИМЕНТНЫЙ ПОРТФЕЛЬ, ВРЕМЕННЫЕ РЯДЫ, ЛИНЕЙНЫЙ КЛЕТОЧНЫЙ АВТОМАТ, SILHOUETTE ANALYSIS, ELBOW METHOD, KNEEDLE

Keywords: PRODUCT PORTFOLIO, TIME SERIES, LINEAR CELLULAR AUTOMATON, SILHOUETTE ANALYSIS, ELBOW METHOD, KNEEDLE

<http://dx.doi.org/10.21515/1990-4665-211-032>

Введение.

В условиях нарастающей конкуренции на рынке бутилированной воды, монопродуктовая компания, специализирующаяся на продаже и доставке питьевой воды в 19-литровых бутылках, сталкивается с вопросами оперативного управления ассортиментом и оптимизации логистических процессов. Фундаментальным элементом эффективного оперативного управления является прецизионный прогноз ежедневных продаж каждой марки воды, входящей в ассортиментный портфель. Микроуровневый анализ, акцентированный на прогнозировании ежедневного спроса для каждой конкретной марки, позволяет компании оперативно адаптироваться к динамике рыночной конъюнктуры, оптимизировать уровень товарных запасов и минимизировать операционные издержки.

Формирование релевантного ассортиментного портфеля монопродуктовой компании, специализирующейся на питьевой воде в 19-литровых бутылках, предполагает дифференциацию марок по ряду определяющих критериев, включая источник воды (артезианская, родниковая, вода прошедшая многоступенчатую фильтрацию), уровень минерализации, органолептические свойства и ценовой сегмент [1] Каждая марка позиционируется в отношении определенного сегмента целевой аудитории, определяя, таким образом, параметры спроса и динамику продаж. Оперативный прогноз ежедневных продаж отдельной позиции ассортиментного портфеля монопродуктовой торговой компании влияет не только на ежедневную тактику практической деятельности, но и на общую стратегию развития бизнеса, включая вопросы оптимизации цепочки поставок продукта от производителя, до конечного потребителя.

Обоснование актуальности исследования и обозначение проблемы.

Оперативный прогноз ежедневных продаж марок питьевой воды 19л. является критически важным компонентом эффективного управления ассортиментом и операционной деятельностью монопродуктовой торговой компании. Интеграция современных методов прогнозирования, учет экзогенных факторов и непрерывный мониторинг позволяют компании оперативно реагировать на изменения рынка, оптимизировать запасы и повысить рентабельность бизнеса, обеспечивая устойчивое конкурентное преимущество.

Методология оперативного прогнозирования ежедневных продаж основываться на углубленном анализе исторических данных, характеристик временного ряда продаж для каждой марки воды, выявлении его специфических свойств [2,3].

Выбор оптимального метода прогнозирования в классике базируется на статистическом анализе временных рядов, включая выявление таких составляющих, как тренд, сезонность и автокорреляция. Для повышения предиктивной точности рекомендуется интеграция различных методов прогнозирования. Опираясь на выводы о наличии специфических свойств в структуре временных рядов объемов продаж питьевой бутилированной воды [2], таких как наличие долговременной памяти, персистентности, в настоящей работе предлагается в качестве прогнозного инструментария использовать информационную систему прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS [4], разработанную на основе авторской методологии (алгоритм CATSF) [5].

Таким образом, прогнозирование объемов реализации конкретных брендов воды представляет собой критически важный компонент результативного управления структурой ассортимента. В свою очередь, остается актуальным вопрос внедрения передовых прогностических

методологий для оптимизации ассортиментной матрицы питьевой воды в 19-литровой таре. В частности, методы машинного обучения позволяют существенно повысить точность прогнозных оценок, что способствует принятию более обоснованных управленческих решений в сфере формирования ассортиментного портфеля, логистических операций и маркетинговых стратегий.

Методы и результаты исследования.

Авторская «Информационная система прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS» [4] реализует алгоритм CATSF, представляющий собой новый подход, основанный на интеграции методов машинного обучения, для реализации автоматизированного подхода к обучению линейной клеточно-автоматной модели. Научная новизна заключается в алгоритмизации и таким образом, как результат, автоматизации выбора оптимального числа термов (или "цветов") в ЛКА, расширяя его инструментарий. Данный подход позволяет избежать субъективных оценок экспертов и снизить ошибку прогнозирования благодаря применению машинного обучения для определения оптимального количества термов в процессе обучения ЛКА. Это способствует повышению объективности и точности моделирования, особенно в части, где экспертная оценка может быть затруднена или предвзята.

Разработанная информационная система прогнозирования динамики временных рядов, основанная на клеточном автомате с автоматизированным обучением (CAFTS) [4], обладает ключевым достоинством – универсальностью. Эта характеристика позволяет специалистам применять ее в различных областях исследований, где потребуется реализовать процедуру прогнозирования на основе анализа временных рядов данных. Система поддерживает работу с данными,

представленными в форматах .xlsx и .csv, которые предварительно подготавливаются пользователем.

Практическое применение и валидация информационной системы, включающей набор алгоритмов машинного обучения, как обучающих прогностическую модель на базе линейного клеточного автомата, осуществлены с использованием ежедневных временных рядов (ВР) объемов реализации бутилированной воды «Ажек». Данные ВР характеризуются фрактальной структурой, детерминированностью и устойчивостью к трендам, что детально изучено и описано в публикации [2].

В рамках данного инструментального средства реализована методология, включающая три принципиальных этапа. Продемонстрируем работу этих этапов на примере модельного ежедневного ВР объемов продаж бутилированной воды марки «Ажек» за период 07.08.2023 по 15.07.2024г.

	A	B	C	D
1	date	values		
2	07.08.2023	150		
3	08.08.2023	55		
4	09.08.2023	235		
5	10.08.2023	291		
6	11.08.2023	279		
7	12.08.2023	359		
8	13.08.2023	1		
9	14.08.2023	141		
10	15.08.2023	453		
11	16.08.2023	251		
12	17.08.2023	400		

Рисунок 1 – Пример подготовки входных данных (модельный фрагмент) для информационной системы прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS.

На первом этапе предварительной обработки ВР, при необходимости, в случае если не был проведен предпрогнозный анализ ВР

аналитик имеет возможность выявить наличие фрактальной структуры, детерминированности и устойчивости к трендам исследуемого ВР, как аналитически, так и графически. Расчет значения показателя Херста дает возможность в рамках использования информационной системы провести данную оценку. Причем, использовать несколько методик оценки значения показателя Херста, на основе регрессионного анализа и метода LSTSQ (наименьших квадратов).

В результате анализа (рисунок 2 и рисунок 3) установлено, что величина индекса Херста варьируется, демонстрируя значения 0.57 и 0.62, в зависимости от примененной методологии расчета, при этом разница между значениями (ΔH) составляет 0.05 [6,7]. На основании классификации цветowych шумов и величины показателя Херста, проанализированный временной ряд демонстрирует характеристики, сопоставимые с серым шумом, при этом вычисленное значение показателя приближается к параметрам, свойственным черному шуму [8]. Таким образом, можно сделать вывод о присущих ВР свойствах трендоустойчивости, фрактальности и наличия долговременной памяти [6].



Рисунок 2 – Графическое представление результатов расчетной процедуры оценки показателя Херста модельного ежедневного ВР объемов продаж бутилированной воды марки «Ажек» за период 07.08.2023 по 15.07.2024г. на основе метода наименьших квадратов.

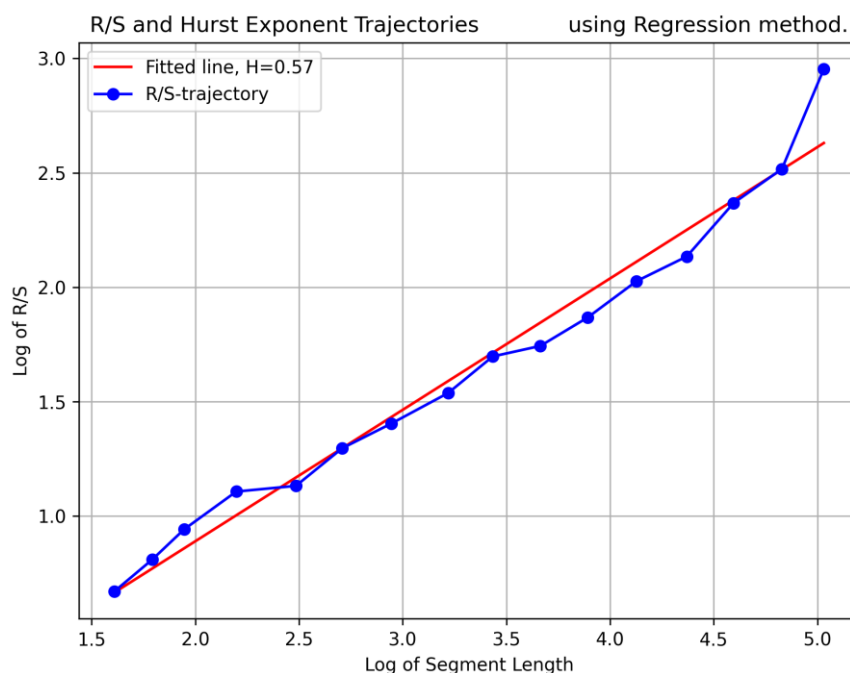


Рисунок 3 – Графическое представление результатов расчетной процедуры оценки показателя Херста модельного ежедневного ВР объемов

продаж бутилированной воды марки «Ажек» за период 07.08.2023 по 15.07.2024г. на основе метода регрессионного анализа.

На втором этапе предварительной обработки значений ВР и процесса автоматизации процедуры кластеризации элементы ВР, на базе алгоритмов машинного обучения для оптимизации данных, делятся на группы (кластеры), каждой из которых сопоставляется уникальный идентификатор (цвет). Данная процедура реализована на основе двух методов - силуэтный анализ (silhouette analysis) и метод локтя (elbow method), дополненный алгоритмом Kneedle.

В разработанной авторской информационной системе [4] предусмотрена возможность графического представления реализации метода силуэтного анализа для выделения кластеров в значениях ВР, причем сверху количество кластеров не ограничено. Рисунок 4 демонстрирует результат кластеризации значений ВР объемов продаж бутилированной воды марки «Ажек» за период с 07.08.2023 по 15.07.2024г. на 3 кластера в 2D-пространстве.

The silhouette plot for the various clusterThe visualization of the clustered data.

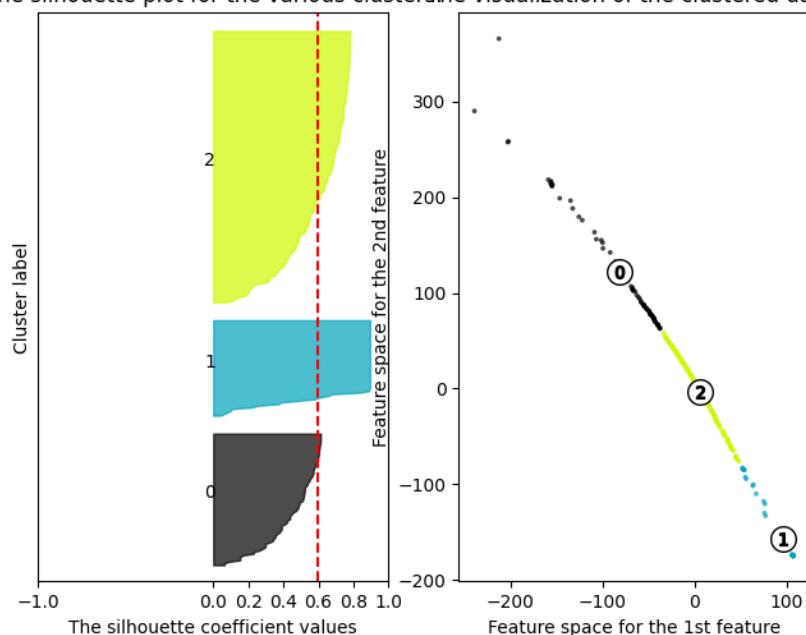


Рисунок 4 – Графическое представление промежуточных результатов силуэтного анализа модельного ежедневного ВР объемов продаж

бутилированной воды марки «Ажек» за период с 07.08.2023 по 15.07.2024г. на три кластера.

Отметим, что приведение матрицы различий к двумерному пространству способствует визуализации фундаментальной структуры временного ряда. Таким образом, аналитик получает возможность визуального мониторинга распределения данных и выявления кластерных тенденций, что обеспечивает интуитивное понимание динамики временного ряда [5]. Классический вариант линейного клеточного автомата базируется именно на 3-х цветной раскраске или в нашем случае на выделении трех кластеров интервальных значений представлено в виде гистограммы на рисунке 5.

Визуализация структуры данных в 2D-пространстве на рисунке 5 позволяет оценить неравномерность количества данных в трех кластерах, что подтверждается интервальными данными: значения, принадлежащие кластеру А находятся в интервале значений [10.00; 108.00]; значения, принадлежащие кластеру В – [116.00; 273.00]; значения, принадлежащие кластеру С – [279.00; 629.00].

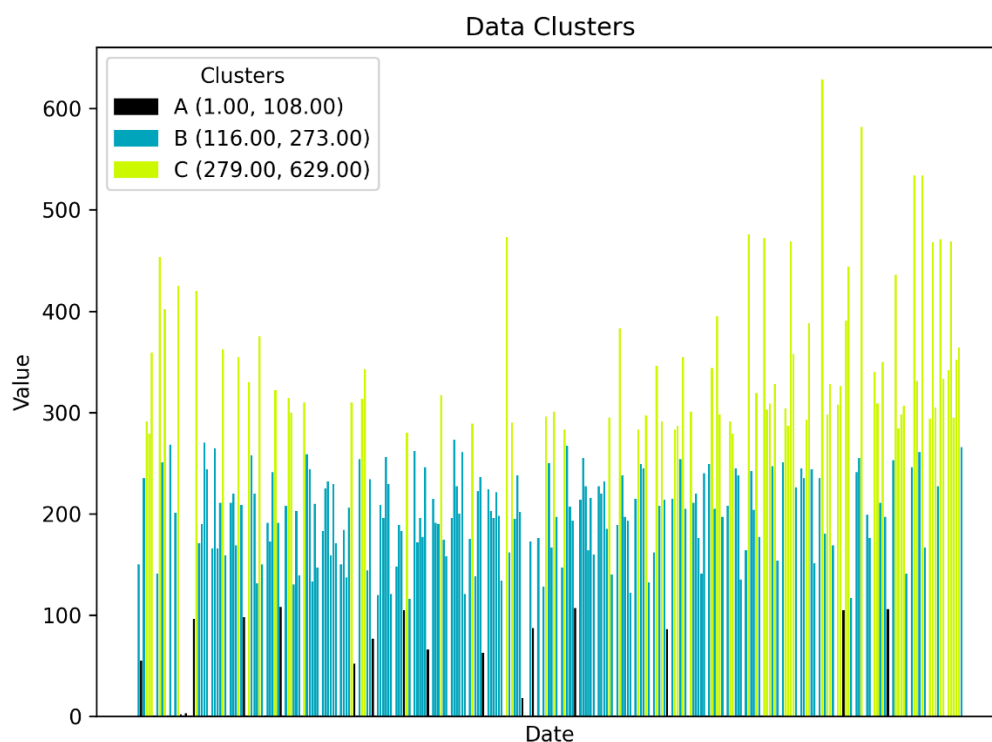


Рисунок 5 – Гистограмма, демонстрирующая работу алгоритмов кластеризации (silhouette analysis) и (elbow method+Kneedle): модельный ВР ежедневных значений объемов продаж бутилированной питьевой воды 19л. торговой марки «Ажек» за период 07.08.2023 по 15.07.2024г. с маркировкой данных в виде категориальной метки – цвета, а именно 3 кластера – 5 цвета, соответственно 3 интервала данных.

Комплексное использование методов кластеризации выступает инструментом поддержки принятия решений для оптимальной кластеризации в вопросе принятия решения о том, как наилучшим образом сегментировать данные, основываясь на их внутренней структуре. На выходе MDS используется алгоритм K-means++ для отнесения каждого наблюдения к определенному кластеру. Затем каждому кластеру присваивается уникальная лингвистическая метка и цвет [5].

Рисунок 6 демонстрирует результат силуэтного анализа данных в двух проекциях. Отметим, что геометрические размеры кластеров

достаточно равномерны относительно кластеризации на 3 части (рисунок 4).

The silhouette plot for the various clusterThe visualization of the clustered data.

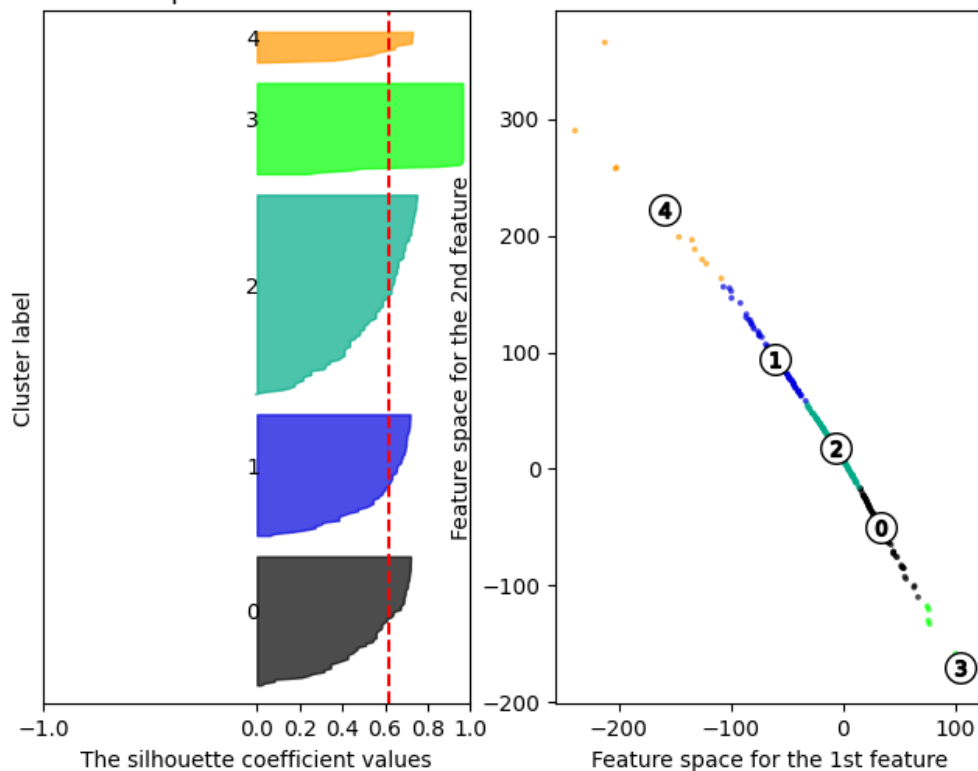


Рисунок 6 – Графическое представление промежуточных результатов силуэтного анализа модельного ежедневного ВР объемов продаж бутилированной воды марки «Ажек» за период с 07.08.2023 по 15.07.2024г. на 5 кластеров.

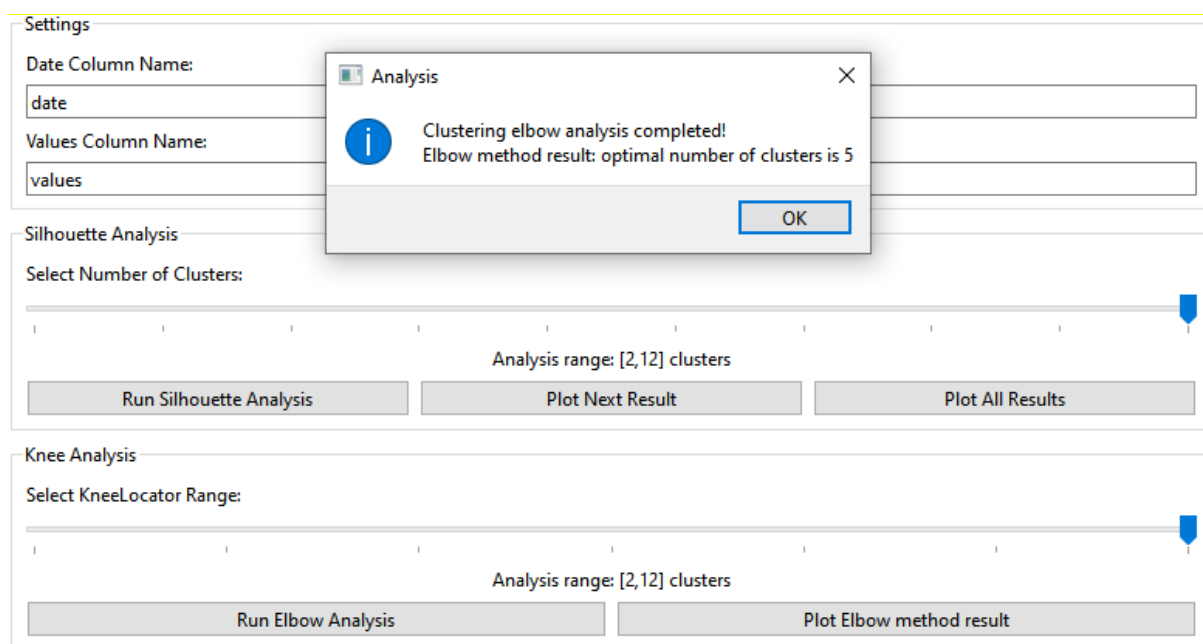


Рисунок 7 – Графическое представление промежуточных результата применения метода локтя, дополненного алгоритмом Kneedle к модельному ежедневному ВР объемов продаж бутилированной воды марки «Ажек» за период с 07.08.2023 по 15.07.2024г.

Результаты анализа данных ВР методом локтя (elbow method), усовершенствованного алгоритмом Kneedle представлены на рисунках 7 и 8. Так, на рисунке 8 представлено графическое обоснование – выделена «точка перегиба» количества кластеров, т.е. метрика SSE – является мерой конечного количества кластеров, т.е. дальнейшее увеличение кластеров и, соответственно, изменение(уменьшение) количества элементов в кластере существенно не влияет на дальнейший результат метрики. По мере роста величины k , сумма квадратов ошибок (SSE) закономерно снижается, что свидетельствует о сокращении дисперсии внутри кластеров [9]. После достижения определенного порога, снижение SSE происходит с меньшей интенсивностью. Данная точка перегиба, часто именуемая "локтем", сигнализирует о достижении оптимального компромисса между числом кластеров и минимизацией внутрикластерной вариабельности [10].

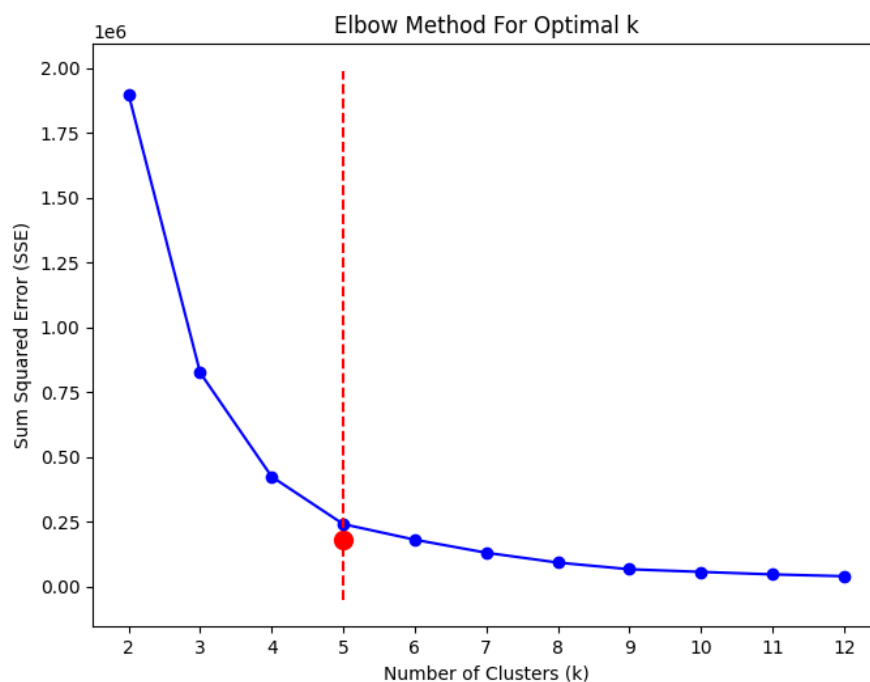


Рисунок 8 – Графическое представление промежуточных результатов применения метода локтя, дополненного алгоритмом Kneedle к модельному ежедневному ВР объемов продаж бутилированной воды марки «Ажек» за период с 07.08.2023 по 15.07.2024г

Ключевым третьим этапом фазой [5], реализованной в комплексной информационной системе выступает сам процесс прогнозирования динамики ВР, основанный на применении алгоритма линейных клеточных автоматов для экстраполяции последующих состояний системы. Данный процесс опирается на анализ временных рядов лингвистических переменных, что соответствует методологии, подробно изложенной в работах [11,12]. Этот подход позволяет моделировать динамику системы, учитывая её лингвистические характеристики и эволюцию во времени. Отметим, что лингвистическая переменная, в данном контексте, представляет собой интервальную величину, определяемую диапазоном между максимальным и минимальным значениями, характерными для соответствующего кластера, что согласуется с принципами нечеткой логики, впервые представленными в работе [13].

Итак, после установления оптимального числа кластеров посредством анализа поддержки принятия решений, дальнейшим этапом является отнесение каждой точки данных к конкретному кластеру в двухмерном пространстве Y . Автоматизированная цветовая кодировка реализуется посредством применения алгоритма кластеризации k -средних++ [14]. K -средних++ представляет собой усовершенствование стандартного метода k -средних, обеспечивающее устойчивую стратегию инициализации, которая распределяет начальные центроиды по всему пространству данных, минимизируя вероятность сходимости к локальным. По завершении кластеризации каждому кластеру назначается уникальный цвет и буквенное обозначение (например, "A", "B", "C" и т.д.), что трансформирует числовые данные временного ряда в дискретное категориальное представление. Данное преобразование имеет существенное значение для последующего прогнозирования, основанного на линейном клеточном автомате, поскольку оно позволяет установить долговременную память посредством лингвистических переменных [12,13].

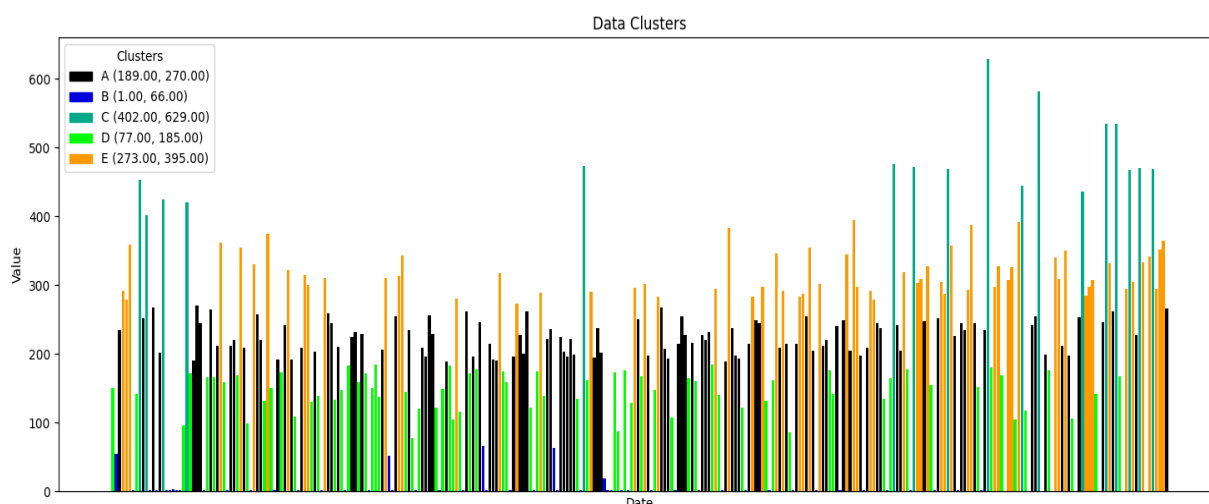


Рисунок 9 – Гистограмма, демонстрирующая работу алгоритмов кластеризации (silhouette analysis) и (elbow method+Kneedle): модельный ВР ежедневных значений объемов продаж бутилированной питьевой воды

19л. торговой марки «Ажек» за период 07.08.2023 по 15.07.2024г. с маркировкой данных в виде категориальной метки – цвета, а именно 5 кластеров – 5 цветов, соответственно 5 интервалов данных.

Результат кластеризации ВР ежедневного объема продаж питьевой воды торговой марки «Ажек» на 5 кластеров, представлен на рисунке 9. Пяти цветная гистограмма демонстрирует «раскраску» ВР в 5 цветов, согласно проведенному кластерному анализу: значения, принадлежащие кластеру В (синий цвет) находятся в интервале значений [10.00; 66.00]; кластер D (ярко зеленый) – [77.00; 185.00]; кластер А (черный) – [189.00; 270.00]; С (зеленый) – [402.00; 629.00]; Е (оранжевый) – [273.00; 395.00]. Отметим, в данном случае, факт достаточной равномерности интервальных данных кластеров А, В, D, Е. Кластер С имеет наименьшее количество значений в интервале, который превосходит в 2 раза по «длине» все 4 кластера А, В, D, Е. Т.е. в этот интервал включены такие редкие события больших объемов отгрузок, которые связаны с определенными событийными составляющими. Например, снабжение питьевой водой мероприятий типа «Формула 1». Подготовка данных для перехода к завершающей фазе, фазе прогнозирования с помощью линейного клеточного автомата продемонстрирована на рисунке 10 в виде фрагмента в формате xls.

	A	B	C	D	E	F
1		date	values	cluster	labels	labels with interval
2	0	2023-08-07 00:00:00	150	4	E	E (77.00, 184.00)
3	1	2023-08-08 00:00:00	55	1	B	B (1.00, 66.00)
4	2	2023-08-09 00:00:00	235	0	A	A (185.00, 270.00)
5	3	2023-08-10 00:00:00	291	3	D	D (273.00, 395.00)
6	4	2023-08-11 00:00:00	279	3	D	D (273.00, 395.00)
7	5	2023-08-12 00:00:00	359	3	D	D (273.00, 395.00)

Рисунок 10 – Фрагмент результата кластеризации – перевод числовых данных ВР в лингвистические переменные с интервальной

принадлежностью для перехода к фазе прогнозирования с помощью линейного клеточного автомата в формате `xlsx`.

Существование и «глубина» долговременной зависимости во временном ряду формируют основу для функционирования прогностической модели, использующей линейный клеточный автомат [5,12]. Этот факт в использовании математического инструментария ЛКА является ключевым для разработки точных прогнозов. Информационная система прогнозирования динамики временных рядов, основанная на клеточном автомате с автоматизированным обучением (CAFTS) [4], решает задачу выявления и количественной оценки глубины памяти временного ряда, представляющей собой фрактальную характеристику. Реализация данной задачи осуществляется посредством алгоритма, детально описанного в работе [5]. В результате проведенной процедуры верификации ВР ежедневного объема продаж питьевой воды торговой марки «Ажек» определена глубина памяти $L^* = 10$. Значение 10 соответствует приблизительно 2 рабочим неделям, т.к. доставка питьевой воды клиентам осуществляется 5-6 дней в неделю. Опираясь на факт наличия долговременной памяти и глубину $L^* = 10$ проведена валидация линейно-клеточной автоматной модели на исследуемом ВР ежедневного объема продаж питьевой воды торговой марки «Ажек». Фрагмент результативной таблицы валидации модели представлен на рисунке 11. Представим подробное описание процедуры валидации: начиная с первого значения временного ряда выделяется 10 элементов, количество элементов определяется на основе глубины памяти ВР $L^* = 10$, далее, согласно значению функции принадлежности, определяется прогнозное нечеткое терм множество. Например, 11-ый элемент ВР определяет прогнозное нечеткое множество $\{(A; 16.0\%); (B; 1.0\%); (C; 56.0\%); (D; 25.0\%); (E; 2.0\%)\}$, т.е. наиболее вероятное прогнозное значение принадлежит терму

С, т.е. соответствует интервалу [402.00; 629.00] и зеленому цвету в гистограмме на рисунке 9. Таким образом, прогноз (в нашем случае это терм-множество С) и переход в вероятностное состояние С совпадают. Следующий 12-ый элемент ВР определяет прогнозное нечеткое множество $\{(A; 28.0\%); (B; 40.0\%); (C; 20.0\%); (D; 11.0\%); (E; 0.0\%)\}$, т.е. наиболее вероятное прогнозное значение принадлежит терму В, т.е. соответствует интервалу [1.00; 66.00] и синему цвету в гистограмме на рисунке 9. Таким образом, так же как для элемента 12 прогноз элемента 13 (в нашем случае это терм-множество В) и переход в вероятностное состояние В совпадают. Вариант не совпадения терм-множества и перехода в вероятностное состояние считаем ошибкой прогноза.

Ошибка прогноза вычисляется в процентном выражении, как процент «не угаданных» лингвистических переменных в общем количестве значений исследуемого временного ряда.

Средняя ошибка прогноза ВР ежедневного объема продаж питьевой воды торговой марки «Ажек» за период за период 07.08.2023 по 15.07.2024г., согласно расчету по формуле $ME = \frac{1}{N} \sum_{t=1}^N |Z(t) - \tilde{Z}(t)|$, составила 0,17. Стандартное отклонение для этого ВР

$SD = \sqrt{\frac{1}{N} \sum_{t=1}^N (\tilde{Z}(t) - ME)^2}$ соответственно равно 0,11.

	A	B	C	D	E	F	G	H
	№	Прогноз	Конфигурация	Переход в состоянии	Сумма значений функции принадлежности	Значение функции принадлежности	Прогнозное нечеткое терм множество	Да/нет
1								
2	1	C	ADACCCBDEAEB	A	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/2+20/42=0.48	0.16	{(A;16.0%.2f); (B;1.0%.2f); (C;56.0%.2f); (D;25.0%.2f); (E;2.0%.2f)}	Нет
3	1	C	ADACCCBDEAEB	B	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/2+1/42=0.02	0.01	{(A;16.0%.2f); (B;1.0%.2f); (C;56.0%.2f); (D;25.0%.2f); (E;2.0%.2f)}	Нет
4	1	C	ADACCCBDEAEB	C	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/1+1/2+8/42=1.69	0.56	{(A;16.0%.2f); (B;1.0%.2f); (C;56.0%.2f); (D;25.0%.2f); (E;2.0%.2f)}	Да
5	1	C	ADACCCBDEAEB	D	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/2+11/42=0.76	0.25	{(A;16.0%.2f); (B;1.0%.2f); (C;56.0%.2f); (D;25.0%.2f); (E;2.0%.2f)}	Нет
6	1	C	ADACCCBDEAEB	E	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/2+2/42=0.05	0.02	{(A;16.0%.2f); (B;1.0%.2f); (C;56.0%.2f); (D;25.0%.2f); (E;2.0%.2f)}	Нет
7	2	B	DACCCBDEAEB	A	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+4/8+23/65=0.85	0.28	{(A;28.0%.2f); (B;40.0%.2f); (C;20.0%.2f); (D;11.0%.2f); (E;0.0%.2f)}	Нет
8	2	B	DACCCBDEAEB	B	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/1+1/8+5/65=1.20	0.40	{(A;28.0%.2f); (B;40.0%.2f); (C;20.0%.2f); (D;11.0%.2f); (E;0.0%.2f)}	Да
9	2	B	DACCCBDEAEB	C	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+3/8+15/65=0.61	0.20	{(A;28.0%.2f); (B;40.0%.2f); (C;20.0%.2f); (D;11.0%.2f); (E;0.0%.2f)}	Нет
10	2	B	DACCCBDEAEB	D	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/8+22/65=0.34	0.11	{(A;28.0%.2f); (B;40.0%.2f); (C;20.0%.2f); (D;11.0%.2f); (E;0.0%.2f)}	Нет
11	2	B	DACCCBDEAEB	E	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/8+0/65=0.00	0.00	{(A;28.0%.2f); (B;40.0%.2f); (C;20.0%.2f); (D;11.0%.2f); (E;0.0%.2f)}	Нет
12	3	A	ACCCBDEAEB	A	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/1+1/5+20/42=1.68	0.56	{(A;56.0%.2f); (B;7.0%.2f); (C;6.0%.2f); (D;29.0%.2f); (E;2.0%.2f)}	Да
13	3	A	ACCCBDEAEB	B	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/5+1/42=0.22	0.07	{(A;56.0%.2f); (B;7.0%.2f); (C;6.0%.2f); (D;29.0%.2f); (E;2.0%.2f)}	Нет
14	3	A	ACCCBDEAEB	C	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/5+8/42=0.19	0.06	{(A;56.0%.2f); (B;7.0%.2f); (C;6.0%.2f); (D;29.0%.2f); (E;2.0%.2f)}	Нет
15	3	A	ACCCBDEAEB	D	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+3/5+11/42=0.86	0.29	{(A;56.0%.2f); (B;7.0%.2f); (C;6.0%.2f); (D;29.0%.2f); (E;2.0%.2f)}	Нет
16	3	A	ACCCBDEAEB	E	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/5+2/42=0.05	0.02	{(A;56.0%.2f); (B;7.0%.2f); (C;6.0%.2f); (D;29.0%.2f); (E;2.0%.2f)}	Нет
17	4	E	CCCBDEAEB	A	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+9/20+42/108=0.84	0.28	{(A;28.0%.2f); (B;2.0%.2f); (C;22.0%.2f); (D;8.0%.2f); (E;40.0%.2f)}	Нет
18	4	E	CCCBDEAEB	B	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/20+8/108=0.07	0.02	{(A;28.0%.2f); (B;2.0%.2f); (C;22.0%.2f); (D;8.0%.2f); (E;40.0%.2f)}	Нет
19	4	E	CCCBDEAEB	C	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/8+20+27/108=0.65	0.22	{(A;28.0%.2f); (B;2.0%.2f); (C;22.0%.2f); (D;8.0%.2f); (E;40.0%.2f)}	Нет
20	4	E	CCCBDEAEB	D	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/20+25/108=0.23	0.08	{(A;28.0%.2f); (B;2.0%.2f); (C;22.0%.2f); (D;8.0%.2f); (E;40.0%.2f)}	Нет
21	4	E	CCCBDEAEB	E	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/3+20+6/108=1.21	0.40	{(A;28.0%.2f); (B;2.0%.2f); (C;22.0%.2f); (D;8.0%.2f); (E;40.0%.2f)}	Да

Рисунок 11 – Фрагмент таблицы, демонстрирующей процесс валидации линейно-клеточной автоматной модели на исследуемом ВР ежедневного объема продаж питьевой воды торговой марки «Ажек» за период с 07.08.2023 по 15.07.2024г.

	A	B	C	D	E	F	G
	№	Прогноз	Конфигурация	Переход в состоянии	Сумма значений функции принадлежности	Значение функции принадлежности	Прогнозное нечеткое терм множество
1							
2	1	D	DACDDDDDA	A	0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/6+9/26+35/105=0.85	0.28	{(A;28%; (B;14%); (C;5%); (D;30%); (E;22%)}
3	1	D	DACDDDDDA	B	0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/6+4/26+10/105=0.42	0.14	{(A;28%; (B;14%); (C;5%); (D;30%); (E;22%)}
4	1	D	DACDDDDDA	C	0/1+0/1+0/1+0/1+0/1+0/1+0/1+0/6+2/26+8/105=0.15	0.05	{(A;28%; (B;14%); (C;5%); (D;30%); (E;22%)}
5	1	D	DACDDDDDA	D	0/1+0/1+0/1+0/1+0/1+0/1+0/1+3/6+5/26+23/105=0.91	0.30	{(A;28%; (B;14%); (C;5%); (D;30%); (E;22%)}
6	1	D	DACDDDDDA	E	0/1+0/1+0/1+0/1+0/1+0/1+0/1+1/6+6/26+29/105=0.67	0.22	{(A;28%; (B;14%); (C;5%); (D;30%); (E;22%)}
7	2	A	ACDDDDDA	A	0/1+0/1+0/1+0/1+0/1+0/1+2/3+3/5+8/23+27/65=2.03	0.51	{(A;51%; (B;2%); (C;2%); (D;21%); (E;25%)}
8	2	A	ACDDDDDA	B	0/1+0/1+0/1+0/1+0/1+0/1+0/3+0/5+1/23+2/65=0.07	0.02	{(A;51%; (B;2%); (C;2%); (D;21%); (E;25%)}
9	2	A	ACDDDDDA	C	0/1+0/1+0/1+0/1+0/1+0/1+0/3+0/5+0/23+4/65=0.06	0.02	{(A;51%; (B;2%); (C;2%); (D;21%); (E;25%)}
10	2	A	ACDDDDDA	D	0/1+0/1+0/1+0/1+0/1+0/1+0/3+1/5+8/23+18/65=0.82	0.21	{(A;51%; (B;2%); (C;2%); (D;21%); (E;25%)}
11	2	A	ACDDDDDA	E	0/1+0/1+0/1+0/1+0/1+0/1+1/3+1/5+6/23+14/65=1.01	0.25	{(A;51%; (B;2%); (C;2%); (D;21%); (E;25%)}
12	3	E	CDDDDDDA	A	0/1+0/1+0/1+0/1+0/1+0/2+0/3+3/8+9/26+35/105=1.05	0.21	{(A;21%; (B;24%); (C;3%); (D;11%); (E;41%)}
13	3	E	CDDDDDDA	B	0/1+0/1+0/1+0/1+0/1+1/2+1/3+1/8+4/26+10/105=1.21	0.24	{(A;21%; (B;24%); (C;3%); (D;11%); (E;41%)}
14	3	E	CDDDDDDA	C	0/1+0/1+0/1+0/1+0/1+0/2+0/3+0/8+2/26+8/105=0.15	0.03	{(A;21%; (B;24%); (C;3%); (D;11%); (E;41%)}
15	3	E	CDDDDDDA	D	0/1+0/1+0/1+0/1+0/1+0/2+0/3+1/8+5/26+23/105=0.54	0.11	{(A;21%; (B;24%); (C;3%); (D;11%); (E;41%)}
16	3	E	CDDDDDDA	E	0/1+0/1+0/1+0/1+0/1+1/2+2/3+3/8+6/26+29/105=2.05	0.41	{(A;21%; (B;24%); (C;3%); (D;11%); (E;41%)}
17	4	B	DDDDDDAA	A	0/1+0/1+0/1+0/1+1/1+1/2+1/3+1/6+8/29+20/69=2.57	0.43	{(A;43%; (B;45%); (C;1%); (D;8%); (E;3%)} {(A;43%); (B;45%); (C;1%); (D;8%); (E;3%)}
18	4	B	DDDDDDAA	B	0/1+0/1+0/1+0/1+0/1+1/2+2/3+4/6+15/29+26/69=2.73	0.45	{(A;43%; (B;45%); (C;1%); (D;8%); (E;3%)} {(A;43%); (B;45%); (C;1%); (D;8%); (E;3%)}
19	4	B	DDDDDDAA	C	0/1+0/1+0/1+0/1+0/1+0/2+0/3+0/6+0/29+3/69=0.04	0.01	{(A;43%; (B;45%); (C;1%); (D;8%); (E;3%)} {(A;43%); (B;45%); (C;1%); (D;8%); (E;3%)}
20	4	B	DDDDDDAA	D	0/1+0/1+0/1+0/1+0/1+0/2+0/3+1/6+4/29+14/69=0.51	0.08	{(A;43%; (B;45%); (C;1%); (D;8%); (E;3%)} {(A;43%); (B;45%); (C;1%); (D;8%); (E;3%)}
21	4	B	DDDDDDAA	E	0/1+0/1+0/1+0/1+0/1+0/2+0/3+0/6+2/29+6/69=0.16	0.03	{(A;43%; (B;45%); (C;1%); (D;8%); (E;3%)} {(A;43%); (B;45%); (C;1%); (D;8%); (E;3%)}
22	5	A	DCDDDAEB	A	0/1+0/1+0/1+1/1+2/2+3/4+6/15+13/26+21/49=4.08	0.68	{(A;68%; (B;3%); (C;3%); (D;9%); (E;17%)}
23	5	A	DCDDDAEB	B	0/1+0/1+0/1+0/1+0/1+0/2+0/4+0/15+0/26+9/49=0.18	0.03	{(A;68%; (B;3%); (C;3%); (D;9%); (E;17%)}
24	5	A	DCDDDAEB	C	0/1+0/1+0/1+0/1+0/1+0/2+0/4+1/15+2/26+2/49=0.18	0.03	{(A;68%; (B;3%); (C;3%); (D;9%); (E;17%)}
25	5	A	DCDDDAEB	D	0/1+0/1+0/1+0/1+0/1+0/2+1/4+2/15+2/26+3/49=0.52	0.09	{(A;68%; (B;3%); (C;3%); (D;9%); (E;17%)}
26	5	A	DCDDDAEB	E	0/1+0/1+0/1+0/1+0/1+0/2+0/4+6/15+9/26+14/49=1.03	0.17	{(A;68%; (B;3%); (C;3%); (D;9%); (E;17%)}
27	6	D	DDDDDAEBA	A	0/1+0/1+0/1+0/1+0/2+0/3+2/6+3/13+6/21+35/105=1.18	0.20	{(A;20%; (B;2%); (C;12%); (D;57%); (E;9%)} {(A;20%); (B;2%); (C;12%); (D;57%); (E;9%)}

Рисунок 12 – Фрагмент таблицы, демонстрирующей процесс прогноза на 6 будущих значений линейно-клеточной автоматной модели для

исследуемого ВР ежедневного объема продаж питьевой воды торговой марки «Ажек» за период за период 07.08.2023 по 15.07.2024г.

Опираясь на выявленный факт наличия и выявленного значения глубины памяти ВР, прогнозная модель ЛКА позволяет вычислить прогноз «вперед» на 10 значений. Рисунок 12 демонстрирует фрагмент процесса прогноза. Итогом прогноза выступает следующая последовательность интервальных термов: D, A, E, B, A, D, D, A, D, A.

Заключение.

Таблица 1. Статистические метрики оценки качества результатов работы классической и комплексной прогнозной модели ЛКА.

№	Метод	Количество кластеров	Глубина памяти	Средняя ошибка прогноза (ME)	Стандартное отклонение прогнозных значений (SD)
1	Классический ЛКА	3	8	0,19	0,13
2	Комплекс ЛКА с машинным обучением, (оптимальное количество кластеров выявлено методом локтя)	5	10	0,16	0,09

Сравнительный анализ результатов работы классической и комплексной прогнозной модели ЛКА представлен в таблице 1 в виде значений статистических метрик. Отметим факт повышения качества прогноза при использовании комплекса методов ЛКА и машинного обучения. Промежуточным, но значимым для повышения показателя качества прогноза результатом является рассчитанное оптимальное количество кластеров, определенное на основе применения методов машинного обучения k-means. Отметим выявленный факт неявного улучшения качества прогноза комплексной модели за счет алгоритмически рассчитанного количества кластеров, которое равно 5, в отличие от

классических трех. Так, если в классической модели выявлена глубина памяти равная 8 и прогноз соответственно может быть рассчитан на 8 лингвистических значений вперед в виде укрупненных интервалов, т.е. значения ВР принадлежащие кластеру А находятся в интервале значений [10.00; 108.00]; значения, принадлежащие кластеру В – [116.00; 273.00]; значения, принадлежащие кластеру С – [279.00; 629.00], то при использовании гибридной модели увеличивается как точность прогноза, согласно значениям метрик Таблицы 1, так такие показатели как количество кластеров, глубина памяти ВР, а также соответственно количеству кластеров уменьшаются интервалы термов: В – [10.00; 66.00]; D – [77.00; 185.00]; А – [189.00; 270.00]; С – [402.00; 629.00]; Е – [273.00; 395.00]. Соответственно, чем меньше интервальное значение терма, тем точнее выявленный прогноз. Отдельно отметим, что в прогнозной результативной выборке гибридного метода отсутствует терм С, что ожидаемо, т.к. данный интервал значений хоть и имеет наибольшую разность предельных значений, но плотность значений в нем наименьшая, а сами значения отражают событийные составляющие, например, снабжение питьевой водой мероприятия «Формула 1». Тем не менее подтверждается факт точного прогноза реверсов (подъемов и спадов) в значениях базового ряда при использовании как классической, так и гибридной модели ЛКА [15], что в свою очередь определяет перспективу использования авторского гибридного подхода для, например, таких временных рядов, значения которых зависят от наличия сезонной составляющей.

Таким образом, предложенный автором [5] гибридный метод не искажает результатов классического, а значительно повышает точность прогноза, работая в отличие от нейро-сетевых прогнозных технологий на коротких временных рядах.

Список использованных источников

1. Попова, М. И. Методы многокритериальной оптимизации в формировании ассортимента монопродуктовой компании / М. И. Попова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2024. – № 202. – С. 151-165. – DOI 10.21515/1990-4665-202-014. – EDN ZYHOJD.
2. Попова, М. И. Предиктивное исследование специфических свойств структуры временных рядов объемов продаж питьевой бутилированной воды / М. И. Попова // Современная экономика: проблемы и решения. – 2025. – № 6 (186). – С. 6-20.
3. Попова, М. И. Декомпозиционный подход к исследованию временных рядов объемов продаж монопродуктового бизнеса сервисного типа / М. И. Попова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2024. – № 203. – С. 440-453. – EDN LMGRWV.
4. Свидетельство о государственной регистрации программы для ЭВМ № 2025611854 Российская Федерация. Информационная система прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS : заявл. 10.01.2025 : опубл. 23.01.2025 / Е. В. Попова, А. Хаммуд, М. И. Попова ; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Кубанский государственный аграрный университет им. И.Т. Трубилина». – EDN HJWCDE.
5. Хаммуд, А. Математическая модель прогнозного инструмента на базе комплекса методов K-means и клеточных автоматов / А. Хаммуд, М. И. Попова, Е. В. Попова // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2025. – № 208. – С. 251-276. – DOI 10.21515/1990-4665-208-024. – EDN XAJZCB.
6. Mandelbrot, B. B., Wallis, J. R. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. Water Resources Research, 1969, 5(5), - С. 967-988.
7. Перепелица, В. А. Фрактальный анализ поведения природных временных рядов / В. А. Перепелица, Е. В. Попова // Современные аспекты экономики. – 2002. – № 9(22). – С. 185-200. – EDN WQITPT.
8. Перепелица, В. А. Комбинированные экономико-математические методы в практике управления налоговым органом / В. А. Перепелица, Д. А. Тамбиева, Е. В. Попова. – Краснодар : Кубанский государственный аграрный университет имени И.Т. Трубилина, 2008. – 130 с. – ISBN 978-5-94672-366-4. – EDN SDWZYZ.
9. Hartigan, J.A. Clustering Algorithms, – 1975. John Wiley & Sons. – 369 pp.
10. Торндайк, Р. Л. Кто принадлежит семье? Психометрика 18, 4 (1953), – С. 267-276.
11. Zadeh, L.A.. Fuzzy sets. Information and Control, 1965, 8(3), – С. 338-353.
12. Перепелица, В.А., Касаева, М.Д. Прогнозирование природного временного ряда на базе модели клеточного автомата / В.А Перепелица, М.Д. Касаева //Современные аспекты экономики. Санкт-Петербург. – 2002. № 9. – С. 209-212].
13. Попова, Е.В. Теория нечетких множеств и клеточных автоматов как инструмент прогноз и адекватного отражения стохастической природы экономических процессов / Е. В. Попова, Н. О. Позднышева, Д. Н. Савинская [и др.] // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2011. – № 67. – С. 173-194. – EDN NEJCPL.
14. Артур, Д., Васильвицкий, С. K-means++: преимущества тщательного выбора начального набора данных. // Материалы восемнадцатого ежегодного симпозиума

АСМ-SIAM по дискретным алгоритмам, SODA 2007, Новый Орлеан, Луизиана, США, 7-9 января 2007 г. – С. 1027-1035.

15. Кумратова, А. М. Применение методов нелинейной динамики и машинного обучения для прогнозирования экономических волатильных процессов / А. М. Кумратова, В. А. Плотников // *π-Economy*. – 2024. – Т. 17, № 3. – С. 81-95. – DOI 10.18721/JE.17306. – EDN BPTKWE.

References

1. Popova, M. I. Metody` mnogokriterial`noj optimizacii v formirovanii assortimenta monoproduktovoj kompanii / M. I. Popova // *Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta*. – 2024. – № 202. – S. 151-165. – DOI 10.21515/1990-4665-202-014. – EDN ZYHOJD.

2. Popova, M. I. Prediktivnoe issledovanie specificheskix svojstv struktury` vremenny`x ryadov ob`emov prodazh pit`evoj butilirovannoj vody` / M. I. Popova // *Sovremennaya e`konomika: problemy` i resheniya*. – 2025. – № 6 (186). – S. 6-20.

3. Popova, M. I. Dekompozicionny`j podxod k issledovaniyu vremenny`x ryadov ob`emov prodazh monoproduktovogo biznesa servisnogo tipa / M. I. Popova // *Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta*. – 2024. – № 203. – S. 440-453. – EDN LMGRWV.

4. Svidetel'stvo o gosudarstvennoj registracii programmy` dlya E`VM № 2025611854 Rossijskaya Federaciya. Informacionnaya sistema prognoza dinamiki vremenny`x ryadov na baze kletchnogo avtomata s avtomatizaciej obucheniya CAFTS : zayavl. 10.01.2025 : opubl. 23.01.2025 / E. V. Popova, A. Xammud, M. I. Popova ; zayavitel` Federal'noe gosudarstvennoe byudzhethoe obrazovatel'noe uchrezhdenie vy`sšhego obrazovaniya «Kubanskij gosudarstvenny`j agrarny`j universitet im. I.T. Trubilina». – EDN HJWCDE.

5. Xammud, A. Matematicheskaya model` prognoznogo instrumenta na baze kompleksa metodov K-means i kletchny`x avtomatov / A. Xammud, M. I. Popova, E. V. Popova // *Politematicheskij setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta*. – 2025. – № 208. – S. 251-276. – DOI 10.21515/1990-4665-208-024. – EDN XAJZCB.

6. Mandelbrot, B. B., Wallis, J. R. Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence. *Water Resources Research*, 1969, 5(5), - C. 967-988.

7. Perepelicz, V. A. Fraktal'ny`j analiz povedeniya prirodny`x vremenny`x ryadov / V. A. Perepelicz, E. V. Popova // *Sovremennye aspekty` e`konomiki*. – 2002. – № 9(22). – S. 185-200. – EDN WQITPT.

8. Perepelicz, V. A. Kombinirovanny`e e`konomiko-matematicheskie metody` v praktike upravleniya nalogovy`m organom / V. A. Perepelicz, D. A. Tambieva, E. V. Popova. – Krasnodar : Kubanskij gosudarstvenny`j agrarny`j universitet imeni I.T. Trubilina, 2008. – 130 s. – ISBN 978-5-94672-366-4. – EDN SDWZYR.

9. Hartigan, J.A. *Clustering Algorithms*, – 1975. John Wiley & Sons. – 369 pp.

10. Torndajk, R. L. Kto prinadlezhit sem'e? *Psixometrika* 18, 4 (1953), – C. 267-276.

11. Zadeh, L.A.. Fuzzy sets. *Information and Control*, 1965, 8(3), – C. 338-353.

12. Perepelicz, V.A., Kasaeva, M.D. Prognozirovanie prirodnogo vremennogo ryada na baze modeli kletchnogo avtomata / V.A Perepelicz, M.D. Kasaeva // *Sovremennye aspekty` e`konomiki. Sankt-Peterburg*. – 2002. № 9. – S. 209-212].

13. Popova, E.V. Teoriya nechetkix mnozhestv i kletchny`x avtomatov kak instrumentarij prognoza i adekvatnogo otrazheniya stoxasticheskoy prirody` e`konomicheskix processov / E. V. Popova, N. O. Pozdny`sheva, D. N. Savinskaya [i dr.] // *Politematicheskij*

setevoj e`lektronny`j nauchny`j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2011. – № 67. – S. 173-194. – EDN NEJCPL.

14. Artur, D., Vasil`viczkij, S. K-means++: preimushhestva tshhatel`nogo vy`bora nachal`nogo nabora danny`x. // Materialy` vosemnadczatogo ezhegodnogo simpoziuma ACM-SIAM po diskretny`m algoritmam, SODA 2007, Novy`j Orlean, Luiziana, SShA, 7-9 yanvarya 2007 g. – S. 1027-1035.

15. Kumratova, A. M. Primenenie metodov nelinejnoj dinamiki i mashinnogo obucheniya dlya prognozirovaniya e`konomicheskix volatil`ny`x processov / A. M. Kumratova, V. A. Plotnikov // π -Economy. – 2024. – T. 17, № 3. – S. 81-95. – DOI 10.18721/JE.17306. – EDN BPTKWE.