

УДК 330.4 JEL C02

5.2.2. Математические, статистические и инструментальные методы экономики (физико-математические науки, экономические науки)

**МАТЕМАТИЧЕСКАЯ МОДЕЛЬ
ПРОГНОЗНОГО ИНСТРУМЕНТА НА БАЗЕ
КОМПЛЕКСА МЕТОДОВ K-MEANS И
КЛЕТОЧНЫХ АВТОМАТОВ**

Хаммуд Али
аспирант

Попова Маргарита Игоревна
ассистент

Попова Елена Витальевна
доктор экономических наук, профессор
ФГБОУ ВО «Кубанский государственный
аграрный университет имени И. Т. Трубилина»,
Краснодар, Россия

В настоящем исследовании авторы развивают инструментарий линейного клеточного автомата, реализуя три основных цели: во-первых, продемонстрировать алгоритм CATSF, как разработанный подход на основе комплексного использования методов машинного обучения с обоснованием выбора оптимального количества «термов» или «цветов» ЛКА, который обладает научной новизной и развивает инструментарий ЛКА, т.к. автоматизация процесса определения количества термов (процесса обучения) ЛКА и обоснование оптимального количества термов позволяет через методы машинного обучения ЛКА исключить субъективное мнение эксперта и уменьшить ошибку прогноза; во-вторых, предоставить экономистам-математикам алгоритмы кластеризации и линейного клеточного автомата в виде комплекса математических моделей в аналитическом формате, т.к. во всех предшествующих научных работах, посвященных использованию модели линейного клеточного автомата, авторы предлагают и используют описательную модель в терминологии теории графов и нечеткой логики; в-третьих, на основе исследования провести сравнение линейного клеточного автомата с клеточным автоматом в его классической постановке. Все цели авторами достигнуты

Ключевые слова: КЛЕТОЧНЫЕ АВТОМАТЫ, МЕТОДЫ КЛАСТЕРИЗАЦИИ, МАШИННОЕ ОБУЧЕНИЕ, ВРЕМЕННЫЕ РЯДЫ

<http://dx.doi.org/10.21515/1990-4665-208-024>

UDC 330.4 JEL C02

5.2.2. Mathematical, statistical and instrumental methods of economics (physical and mathematical sciences, economic sciences)

**MATHEMATICAL MODEL OF
A FORECASTING TOOL BASED ON THE K-
MEANS METHODS AND CELLULAR
AUTOMATA COMPLEX**

Hammoud Ali
graduate student

Popova Margarita Igorevna
assistant

Popova Elena Vitalievna
Doctor of Economics, Professor
Kuban State Agrarian University named after I. T.
Trubilin, Krasnodar, Russia

In this study, the authors develop the toolkit of a linear cellular automaton, realizing three main goals: first, to demonstrate the CATSF algorithm as a developed approach based on the integrated use of machine learning methods to justify the choice of the optimal number of "terms" or "colors" of a linear cellular automaton, which has scientific novelty and develops the toolkit of a linear cellular automaton. because automating the process of determining the number of terms (learning process) of a linear cellular automaton and substantiating the optimal number of terms allows, through the methods of machine learning of a linear cellular automaton, to exclude the subjective opinion of an expert and reduce the forecast error; secondly, to provide mathematical economists with clustering algorithms and linear cellular automaton in the form of a complex of mathematical models, in an analytical format. In previous scientific papers devoted to the use of the linear cellular automaton model, the authors propose and use a descriptive model in the terminology of graph theory and fuzzy logic; thirdly, based on the study, to compare a linear cellular automaton with a cellular automaton in its classical formulation. All goals were achieved by the authors

Keywords: CELLULAR AUTOMATA, CLUSTERING METHODS, MACHINE LEARNING, TIME SERIES

Введение

Вопросы автоматизации обучения и самообучения нейронных сетей, генетических алгоритмов и систем искусственного интеллекта как прогнозных моделей отражающих динамику экономических процессов являются актуальными на современном этапе развития науки и общества. Особое внимание авторами уделяется вопросам алгоритмизации и развития методов машинного обучения в прогнозных моделях. Интеллектуальные методы анализа позволяют исследовать сложные нелинейные особенности, свойственные экономическим процессам с высокой нестабильностью и неопределенностью, с наличием скрытых сложно выявляемых зависимостей. Линейный клеточный автомат (ЛКА) представляет собою такой интеллектуальный инструментарий, который позволяет исследовать динамику временных рядов, обладающих фрактальными свойствами, выявлять долговременную память, и при ее наличии осуществлять прогноз. Именно этому прогнозному инструментарию и вопросам его совершенствования посвящена настоящая работа.

Обоснование актуальности исследования и обозначение проблемы.

История развития и становления метода прогнозирования динамики экономических показателей под названием «Линейный клеточный автомат» началась с публикации российских авторов Перепелицы В.А. и его ученицы Касеовой М.Д. в 2002 году в журнале «Современные аспекты экономики» [3]. Именно в этой статье впервые предложен, уже ставший классическим, метод линейного клеточного автомата (ЛКА), апробированный на ВР данных урожайности пшеницы в Ставропольском крае за период с 1870 по 1996гг. В течение последующих 23 лет этот метод прогноза слабоструктурируемых ВР, неподчиняющихся нормальному

закону распределения, имеющих так называемые «тяжелые хвосты» активно развивался Поповой Е., Тамбиевой Д., Тебуевой Ф., Темировой Л., Леншовой Т., Савинской Д., Замотайловой Д., Шебзуховой М., Болатовой Л., Кошта Л. применительно к различным сферам экономики, таким как промышленность, строительство, страховое дело, казначейские вопросы, туристический бизнес, торговля, финансовые рынки, включая природные процессы и явления [4-11]. Одним из ярких представителей, активно развивающих данный метод прогнозирования, является Кумратова А.М., серия ее работ с соавторами посвящена адаптации и существенному развитию методов нелинейной динамики, включая ЛКА и методы машинного обучения [4,10,11].

Развитие инструментария ЛКА, используя методы машинного обучения является актуальным направлением, способствующим повышению качества принятия решений. В настоящем исследовании авторы развивают инструментарий линейного клеточного автомата, реализуя три основных цели: во-первых, продемонстрировать алгоритм SATSF, как разработанный подход на основе комплексного использования методов машинного обучения с обоснованием выбора оптимального количества «термов» или «цветов» ЛКА, который обладает научной новизной и развивает инструментарий ЛКА, т.к. автоматизация процесса определения количества термов (процесса обучения) ЛКА и обоснование оптимального количества термов позволяет через методы машинного обучения ЛКА исключить субъективное мнение эксперта и уменьшить ошибку прогноза; во-вторых, предоставить экономистам-математикам алгоритмы кластеризации и линейного клеточного автомата в виде комплекса математических моделей, в аналитическом формате, т.к. в предшествующих научных работах, посвященных использованию модели линейного клеточного автомата, авторы предлагают и используют описательную модель в терминологии теории графов и нечеткой логики; в-

третьих, на основе исследования провести сравнение линейного клеточного автомата с клеточным автоматом в его классической постановке.

Методы и результаты исследования

Авторская методология (алгоритм CATSF) уникально сочетает в себе передовые методы кластеризации с принципами клеточных автоматов для сегментирования и прогнозирования на основе временных рядов. Методология включает две основные фазы: (1) фазу предварительной обработки и автоматического «раскрашивания» элементов временного ряда, которая использует машинное обучение для определения оптимального разделения данных на кластеры, каждому из которых присваивается определенный цвет, и (2) основную фазу прогнозирования с использованием линейных клеточных автоматов для прогнозирования будущих состояний на основе временного ряда лингвистических переменных. Лингвистическая переменная отражает интервальное значение в пределах разницы максимального и минимального значения кластера.

На первом этапе данные временных рядов анализируются для выявления естественных группировок или кластеров. Мы используем такие методы, как силуэтный анализ (silhouette analysis) и метод локтя (elbow method), усовершенствованный алгоритмом Kneedle, для определения оптимального количества кластеров, элементам которых соответствует определенный цвет. Кластеры используются для маркировки данных, преобразуют исходные числовые значения в категориальную информацию (метки) или лингвистические (интервальные) значения. На втором этапе алгоритм CATSF использует эти метки для построения

памяти временного ряда и вычисления вероятностей перехода. Представим подробно эти взаимосвязанные этапы.

1. Преобразование данных и построение матрицы различий: исходные данные временных рядов сначала преобразуются в математическое представление, которое фиксирует попарные сравнения между наблюдениями. Это достигается путем построения матрицы сравнений, которая является базовой структурой для последующего анализа.

2. Уменьшение размерности с помощью многомерного масштабирования (MDS): многомерная матрица различий сводится к двумерному представлению. Этот этап не только облегчает визуализацию, но и сохраняет основную информацию о различиях в компактной форме, способствуя эффективной кластеризации.

3. Поддержка принятия решений для оптимальной кластеризации: оптимальное количество кластеров (цветов) определяется с помощью силуэтного анализа (silhouette analysis) и метода локтя (elbow method), усовершенствованного алгоритмом Kneedle. Этот анализ помогает пользователю принять решение о том, как наилучшим образом сегментировать данные, основываясь на их внутренней структуре.

4. Автоматическое окрашивание с помощью кластеризации K-means++: после принятия решения об оптимальном количестве кластеров на выходе MDS используется алгоритм K-means++ для отнесения каждого наблюдения к определенному кластеру. Затем каждому кластеру присваивается уникальная лингвистическая метка и цвет.

5. Прогнозирование клеточных автоматов (CATSF): расширенный набор данных, содержащий как числовые значения, так и категориальные метки, используется для выявления долговременной памяти. Алгоритм CATSF вычисляет вероятности перехода на основе этих

последовательностей и прогнозирует будущие состояния с помощью алгоритма линейного клеточного автомата.

При анализе временных рядов наиболее информативными являются относительные различия между точками данных, а не между их абсолютных значений. Матрица различий является фундаментальным инструментом для их количественной оценки. Используя попарные сравнения, фиксируем внутреннюю структуру временных рядов, которая служит исходным материалом для уменьшения размерности и кластеризации. Это представление особенно полезно, когда последующие шаги требуют понимания относительного расстояния между наблюдениями.

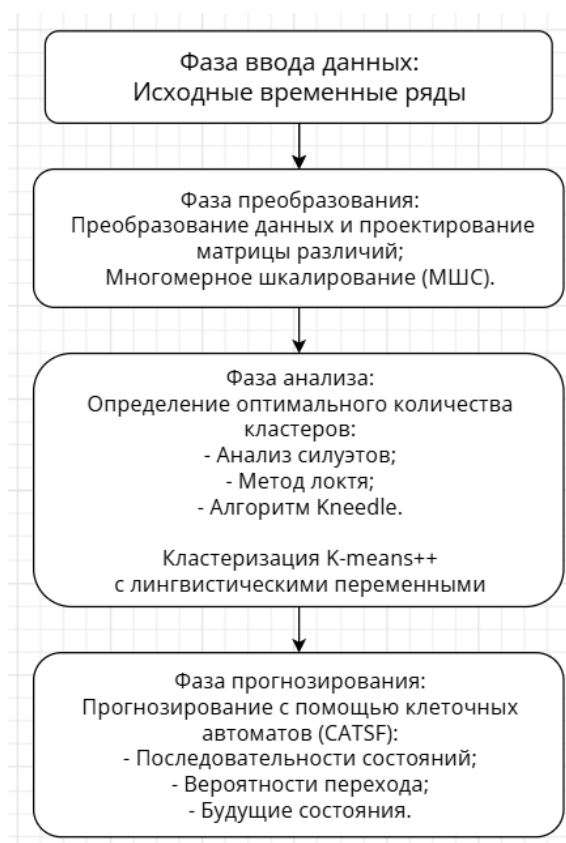


Рисунок 1. Схема фаз реализации алгоритма CATSF

Математическая формулировка:

Рассмотрим одномерный временной ряд, представленный в виде:

$$\mathbf{x} = [x_1, x_2, \dots, x_N],$$

где N – количество наблюдений. Матрица различий D определяется как матрица размерностью $N \times N$, где каждый элемент $D_{i,j}$ измеряет абсолютную разницу между i -м и j -м наблюдениями:

$$D_{i,j} = |x_i - x_j|, \forall i, j \in \{1, 2, \dots, N\}.$$

Этот состав гарантирует, что:

- $D_{i,j} \geq 0$ для всех i и j , что отражает неотрицательный характер абсолютных разностей.
- $D_{i,j} = 0$ тогда и только тогда, когда $x_i = x_j$.
- Матрица D симметрична, т.е. $D_{i,j} = D_{j,i}$.

Роль матрицы различий в процессе прогнозирования: матрица различий D кодирует попарные расстояния между всеми наблюдениями во временном ряду, тем самым фиксируя временную изменчивость данных. Затем эта матрица используется в качестве входных данных для процедуры многомерного масштабирования (MDS), которая направлена на сохранение этих попарных различий в пространстве меньшей размерности [12]. Создавая D , нами закладывается математическая основа для преобразования необработанных данных временных рядов в форму, пригодную для кластеризации и последующего прогнозирования.

После построения матрицы различий следующим шагом является преобразование этого многомерного представления в более понятную форму. Многомерное масштабирование (MDS) — это статистический метод, который преобразует многомерные данные в пространство меньшей размерности, сохраняя структуру попарных расстояний настолько точно, насколько это возможно. В контексте нашей системы прогнозирования MDS используется для проецирования матрицы различий $N \times N$ в

двумерное пространство, тем самым облегчая визуализацию и последующую кластеризацию.

Математическая формулировка: учитывая матрицу различий $\mathbf{D} \in \mathbb{R}^{N \times N}$, MDS ищет набор точек $\mathbf{Y} \in \mathbb{R}^{N \times 2}$ таких, чтобы евклидовы расстояния между этими точками приближались к исходным. Формально пусть $\mathbf{y}_i \in \mathbb{R}^2$ - двумерное представление i -го наблюдения. Целью метода MDS является минимизация так называемой функции напряжения, определяемой как:

$$Stress(\mathbf{Y}) = \sum_{i < j} \frac{(\mathbf{D}_{i,j} - \|\mathbf{y}_i - \mathbf{y}_j\|_2)^2}{\mathbf{D}_{i,j}^2}$$

Здесь $\|\mathbf{y}_i - \mathbf{y}_j\|_2$ представляет евклидово расстояние между вложенными точками $\mathbf{y}_i$ и $\mathbf{y}_j$. Минимизация функции напряжений гарантирует, что конфигурация $\mathbf{Y}$ максимально точно сохранит исходные различия, хотя и в пространстве меньшей размерности.

Роль многомерного масштабирования в процессе прогнозирования: 2D-встраивание, полученное с помощью MDS, служит двум основным целям в рамках предлагаемой методологии: визуализации и кластеризации с дальнейшим анализом.

Визуализация: Сведение матрицы различий к 2D-пространству облегчает визуализацию базовой структуры временного ряда. Исследователи могут визуально отслеживать распределение данных и тенденции кластеризации, получая интуитивное представление о их временной динамике.

Входные данные для кластеризации и анализа: Преобразованные в MDS данные $\mathbf{Y}$ служат входными данными для алгоритмов кластеризации (например, k-means++). Работа в пространстве меньшей размерности не только повышает вычислительную эффективность, но и повышает

надежность процесса кластеризации. Поскольку MDS стремится сохранить исходные попарные расстояния, кластеризация, выполняемая для Y , точно отражает внутренние отношения различий, зафиксированные D .

После процедуры уменьшения размерности двумерное вложение Y представляет внутреннюю структуру данных временного ряда. Чтобы эффективно сегментировать данные на значимые группы, важно определить оптимальное количество кластеров — или, что эквивалентно, оптимальное количество цветов, - которые наилучшим образом отразят лежащие в их основе закономерности. Использование комплекса методов анализа силуэта и метода локтя, дополненного алгоритмом Kneedle позволяют определить и обосновать количество кластеров:

Силуэтный анализ обеспечивает количественную оценку того, насколько каждая точка данных вписывается в назначенный ей кластер по сравнению с другими кластерами [13]. Для каждого наблюдения, y_i принадлежащего кластеру C_m , коэффициент силуэта s_i определяется как:

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)},$$

где: a_i — среднее евклидово расстояние между y_i и всеми другими точками в пределах одного и того же кластера C_m ,

b_i — минимальное среднее евклидово расстояние между y_i и всеми точками ближайшего соседнего кластера.

Общий средний балл по силуэту, рассчитанный как

$$s = \frac{1}{N} \sum_{i=1}^N s_i,$$

обеспечивает оценку сплоченности и разделения кластеров, при этом более высокие значения указывают на более четкую структуру кластеризации. Этот анализ определяет количество кластеров, которое дает наивысший

средний балл по силуэту, предполагая естественную группировку данных [13].

Метод локтя оценивает сумму квадратов ошибок (SSE) для разного количества кластеров k , где SSE определяется как:

$$SSE(k) = \sum_{i=1}^k \sum_{y \in C_i} \|y - \mu_i\|_2^2,$$

где μ_i , — представляющим центр тяжести скопления C_i . По мере увеличения k значение SSE уменьшается, отражая уменьшение внутри кластерной дисперсии. Однако предельное значение SSE уменьшается с меньшей амплитудой после определенного момента. Эта точка перегиба — "колени" — указывает на оптимальный баланс между количеством кластеров и внутри кластерной дисперсией.

Для автоматизации определения этой точки перегиба применяется алгоритм Kneedle [14]. Полином аппроксимирует кривую SSE, и алгоритм определяет точку k^* , в которой упорядоченная вторая производная (кривизна) функции SSE максимальна:

$$k^* = \operatorname{argmax}_k \left(\frac{\partial^2 SSE(k)}{\partial k^2} \right).$$

Достоинство этого метода заключается в том, что он представляет собою объективный критерий для выбора k^* , обеспечивая, таким образом, поддержку принятия решений для определения оптимального количества цветов (кластеров).

Роль метода локтя, дополненный алгоритмом Kneedle в процессе прогнозирования: методы, описанные выше, служат важным инструментом поддержки принятия решений для пользователей. Изучая как баллы силуэта, так и кривую SSE, пользователи могут сделать обоснованный выбор относительно оптимального количества кластеров.

Это решение напрямую влияет на последующий этап автоматической раскраски, на котором выбранное количество кластеров является входной информацией и используется в алгоритмах кластеризации k-means++. В конечном счете, расширенный набор данных — с присвоенными ему метками и цветами — формирует основу для модели прогнозирования временных рядов клеточных автоматов (CATSF).

Поясним последовательно процедуру автоматического «раскрашивания» ЛКА с использованием методов K-means ++.

После определения оптимального количества кластеров с помощью анализа поддержки принятия решений следующим шагом является присвоение каждой точки данных к определенному кластеру в двумерном вложении Y . Автоматическое раскрашивание достигает этого за счет использования алгоритма кластеризации k-means ++. K-means++ улучшает стандартный метод k-средних, предоставляя надежную стратегию инициализации, которая распределяет начальные центроиды по пространству данных. После выполнения кластеризации каждому кластеру присваивается уникальный цвет и буквенная метка (например, "А", "В", "С" и т.д.), преобразующая числовые данные временного ряда в дискретное категориальное представление. Это преобразование важно для последующего прогнозирования на основе линейного клеточного автомата, поскольку позволяет определить долговременную память на основе лингвистических переменных.

Математическая формулировка:

Учитывая двумерное вложение $Y \in \mathbb{R}^{N \times 2}$, k-means++ алгоритм разделяет данные на k кластеров $C_1, C_2, \dots, C_k, \dots$ минимизируя внутрикластерную сумму квадратов ошибок (SSE). Целевая функция определяется как:

$$\min_{C_1, \dots, C_k} \sum_{i=1}^k \sum_{y \in C_i} \|y - \mu_i\|_2^2,$$

где μ_i – обозначает центр тяжести кластера C_i , вычисляемый как:

$$\mu_i = \frac{1}{|C_i|} \sum_{y \in C_i} y.$$

Инициализация k-means++ выбирает первый центроид случайным образом из точек данных. Последующие центроиды выбираются вероятностно, выбирая точки, которые находятся далеко от уже выбранных центроидов. Такой подход максимизирует начальные межцентровые расстояния, тем самым способствуя конвергенции для лучшей конфигурации кластеризации.

Как только кластеры сформированы, каждому наблюдению присваивается кластерный индекс. Затем эти индексы преобразуются в буквенные обозначения, обычно с использованием первых нескольких букв алфавита (например, 'A', 'B', 'C', ...), и сопоставляются определенным цветам с помощью цветовой карты. Этот процесс дает расширенный набор данных, в котором каждая точка данных теперь связана с временным идентификатором, ее исходным значением, категориальной (лингвистической) меткой и соответствующим цветом.

Роль K-means++ в процессе прогнозирования: этап автоматической раскраски переводит численные данные временного ряда в лингвистический(интервальный) для дальнейшего прогнозирования на базе линейного клеточного автомата. Алгоритм эффективно сегментирует временной ряд на дискретные состояния путем преобразования числовых значений в лингвистические с цветовой кодировкой. Такая сегментация не только помогает в визуальной интерпретации, но и облегчает нахождение долговременной памяти временного ряда, что являются неотъемлемой

частью модели CATSF. Категориальные метки, полученные в процессе автоматической раскраски, формируют входные данные для линейного клеточного автомата, где вычисляются вероятности перехода между состояниями, которые используются для прогнозирования будущих значений.

Опишем в аналитической форме прогнозную модель линейного клеточного автомата (CATSF) на базе временных рядов.

Пусть $u_0(t) \in \mathcal{X}$ - наблюдаемое категориальное состояние в момент времени t , где $\mathcal{X}$ - набор конечных состояний.

Эффективная глубина памяти L^* определяется динамически на основе временного ряда лингвистических переменных:

$$L^* = \max\{l \mid C(S[0:l])\} \quad (1)$$

где $S[0:l]$ – последовательность длиной l ,

$C(\cdot)$ – счетчик частоты.

В терминах ЛКА [3]:

l – конфигурация, которая состоит из отрезка временного ряда лингвистических переменных длины l .

$C(\cdot)$ - количество переходов из отрезка длиной l в одно из состояний категориальной переменной.

Рассмотрим практический пример:

Предположим, что результатом процесса кластеризации численные значения временного ряда переведены в лингвистические переменные и такой экспериментальный временной ряд (ЭВР) имеет вид АВАСАВАССАВ.

Далее, используя (1) и (2) рассматриваем варианты возможных переходов или осуществляем поиск глубины памяти.

Таблица 1 – Пример определения глубины памяти временного ряда

Глубина	Переход из	Переход в	Количество переходов	Веса переходов, формула (2)
$l = 1$	A	A	0	$0/5 = 0$
		B	3	$3/5 = 0,6$
		C	2	$2/5 = 0,4$
	B	A	2	$2/2 = 1$
		B	0	$0/2 = 0$
		C	0	$0/2 = 0$
	C	A	2	$2/3 = 0,66$
		B	0	$0/3 = 0$
		C	1	$1/3 = 0,33$
$l = 2$	AB	A	2	$2/2 = 1$
		B	0	$0/2 = 0$
		C	0	$0/2 = 0$
	AC	A	1	$1/2 = 0,5$
		B	0	$0/2 = 0$
		C	1	$1/2 = 0,5$
	BA	A	0	$0/2 = 0$
		B	0	$0/2 = 0$
		C	2	$2/2 = 1$
	CA	A	0	$0/2 = 0$
		B	2	$2/2 = 1$
		C	0	$0/2 = 0$
	CC	A	1	$1/1 = 1$
		B	0	$0/1 = 0$
		C	0	$0/1 = 0$
$l = 3$	ABA	A	0	$0/2 = 0$
		B	0	$0/2 = 0$
		C	2	$2/2 = 1$
	ACA	A	0	$0/1 = 0$
		B	1	$1/1 = 1$
		C	0	$0/1 = 0$
	ACC	A	1	$1/1 = 1$
		B	0	$0/1 = 0$
		C	0	$0/1 = 0$
	BAC	A	1	$1/2 = 0,5$
		B	0	$0/2 = 0$
		C	1	$1/2 = 0,5$
	CAB	A	1	$1/1 = 1$
		B	0	$0/1 = 0$
		C	0	$0/1 = 0$
$l = 4$	ABAC	A	1	$1/2 = 0,5$
		B	0	$0/2 = 0$
		C	1	$1/2 = 0,5$
	BACA	A	0	$0/1 = 0$

		B	1	1/1 = 1
		C	0	0/1 = 0
	BACC	A	1	1/1 = 1
		B	0	0/1 = 0
		C	0	0/1 = 0
$l = 5$	ABACA	A	0	0/1 = 0
		B	1	1/1 = 1
		C	0	0/1 = 0
	ABACC	A	1	1/1 = 1
		B	0	0/1 = 0
		C	0	0/1 = 0

Следует уточнить, что итеративный процесс определения долговременной памяти ВР завершается при соблюдении 2 условий: во-первых количество переходов должно быть формата – 1 переход в 1 состояние, остальные переходы 0 (нулевые); во-вторых, веса переходов имеют те же значения, а именно 1, если есть 1 переход и 0 для других. Демонстрируют соблюдение этих двух условий расчетные значения, находящиеся в окрашенных в серый цвет строках таблицы 1.

В результате проведенной процедуры верификации определена глубина памяти $L^* = \max \{2,3,4,5\} = 5$.

Для времени t (с $t > L^*$) интервал временного ряда, используемый для прогнозирования, имеет вид:

$$S(t) = (u_0(t - L^*), u_0(t - L^* + 1), \dots, u_0(t - 1))$$

Для каждого следующего состояния $x \in \mathcal{X}$ и перехода состояния $S(t)$, пусть:

$$s_i(x; t) = S(t)[i:] \oplus x, \quad \text{для } i = 0, 1, \dots, L^* - 1$$

где $S(t)[i:]$ обозначает переход, начинающееся с i -го элемента. Затем вес перехода или ненормированное значение функции принадлежности перехода в последующее состояние определяется как:

$$W(s_i(x; t)) = \frac{c(S(t)[i:] \oplus x)}{c(S(t)[i:])} \quad (2)$$

Это соотношение количественно определяет вероятность перехода в состояние x , где x – категориальное состояние.

Вычисляем сумму значений функции принадлежности перехода в состояние x путем суммирования взвешенных вкладов всех допустимых переходов:

$$\tilde{P}(x | S(t)) = \sum_{i=0}^{L^*-1} W(S(t)[i:] \oplus x) \quad (3)$$

Нормализация или вычисление значения функции принадлежности показывает вероятностный результат прогноза в виде лингвистической переменной:

$$P(x | S(t)) = \frac{\tilde{P}(x | S(t))}{\sum_{x' \in \mathcal{X}} \tilde{P}(x' | S(t))} \quad (4)$$

Прогноз следующего значения временного ряда в категориальном виде вычисляется следующим образом:

$$\tilde{x}(t) = \operatorname{argmax}_{x \in \mathcal{X}} P(x | S(t)) \quad (5)$$

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

Все вышеописанные шаги построения прогноза входят в процедуру валидации математической прогнозной модели на базе КА. Демонстрация процедуры валидации на примере представлена в Таблице 2.

Таблица 2 – Верификация и валидация прогнозной модели линейного клеточного автомата

№	Прогноз	Конфигурация	Переход в состояние	Сумма значений функции принадлежности Формула (3)	Значение функции принадлежности Формула (4)	Прогнозное нечеткое терм множество	Да/нет
						Формула (5)	
12	А	АССАВ	А	$\frac{C(BA)}{C(B)} + \frac{C(ABA)}{C(AB)} + \frac{C(CABA)}{C(CAB)} +$ $\frac{C(CCABA)}{C(CCAB)} + \frac{C(ACCABA)}{C(ACCAB)} =$ $\frac{2}{2} + \frac{2}{2} + \frac{1}{1} + 0 + 0 = 3$	$\frac{3}{3 + 0 + 0} = 1$	{(A; 100%); (B; 0); (C; 0)}	
			В	$\frac{C(BB)}{C(B)} + \frac{C(ABB)}{C(AB)} + \frac{C(CABB)}{C(CAB)} +$ $\frac{C(CCABB)}{C(CCAB)} + \frac{C(ACCABB)}{C(ACCAB)} =$ $= 0 + 0 + 0 + 0 + 0 = 0$	0		
			С	$\frac{C(BC)}{C(B)} + \frac{C(ABC)}{C(AB)} + \frac{C(CABC)}{C(CAB)} +$ $\frac{C(CCABC)}{C(CCAB)} + \frac{C(ACCABC)}{C(ACCAB)} =$ $= 0 + 0 + 0 + 0 + 0 = 0$	0		
11	В	ВАССА	А	$0 + 0 + 0 + 0 + 0 = 0$	0	{(A; 0); (B; 80%); (C; 20%)}	+
			В	$\frac{3}{5} + \frac{2}{2} + 0 + 0 + 0 =$ $= 1,6$	$1,6/2=0,80$		
			С	$\frac{2}{5} + \frac{0}{2} + 0 + 0 + 0 = 0,4$	$0,4/2=0,20$		
10	А	АВАСС	А	$\frac{2}{3} + 1 + 1 + 1 + 1 = 4,67$	0,93	{(A; 93%); (B; 0); (C; 7%)}	+
			В	$0 + 0 + 0 + 0 + 0 = 0$	0		
			С	$\frac{1}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + 0 = 1,83$	0,07		
9	А	САВАС	А	$\frac{2}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + 0 = 2,17$	0,54	{(A; 54%); (B; 0); (C; 46%)}	-
			В	$0 + 0 + 0 + 0 + 0 = 0$	0		
			С	$\frac{1}{3} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + 1 = 2,83$	0,46		
8	С	АСАВА	А	$0 + 0 + 0 + 0 + 0 = 0$	0	{(A; 0); (B; 20%); (C; 80%)}	+
			В	$\frac{3}{5} + 0 + 0 + 0 + 0 = 0,6$	0,20		
			С	$\frac{2}{5} + 1 + 1 + 0 + 0 = 2,4$	0,80		

6	A	BACAB	A	$1 + 1 + 1 + 0 + 0 = 3$	1	{(A; 100%); (B; 0); (C; 0)}	+
			B	$0 + 0 + 0 + 0 + 0 = 0$	0		
			C	$0 + 0 + 0 + 0 + 0 = 0$	0		
5	B	ABACA	A	$0 + 0 + 0 + 0 + 0 = 0$	0	{(A; 0); (B; 92%); (C; 8%)}	+
			B	$\frac{3}{5} + 1 + 1 + 1 + 1 = 4,6$	0,92		
			C	$\frac{2}{5} + 0 + 0 + 0 + 0 = 0,4$	0,08		

Для пояснений последовательных манипуляций процедуры верификации и валидации с лингвистическими данными ВР пронумеруем значения модельного временного ряда следующим образом:

Таблица 3. Демонстрация модельного ВР для пояснения расчетов верификации и валидации модели ЛКА

	1	2	3	4	5	6	7	8	9	10	11	12
ЭВР	A	B	A	C	A	B	A	C	C	A	B	
Прогноз						B	A	C	A	A	B	A
						92%	100%	80%	54%	93%	80%	100%

Правило расчетов переходов: первая строка таблицы 2 демонстрирует процедуру прогноза на один шаг вперед. Какая лингвистическая переменная выступит прогнозом A, B или C? Опираемся на глубину памяти временного ряда, т.е. в качестве базы прогноза используем пять лингвистических переменных, завершающих временной ряд. Далее, согласно формуле (3), суммируем взвешенные вклады всех допустимых переходов. Отношение переходов $\frac{C(BA)}{C(B)} = \frac{2}{2}$, т.е. в модельном ВР выбираем количество возможных переходов лингвистических переменных, так, всего переходов BA в состояние A, B или C равно 2, или $C(BA) = 2$. Стандартный вопрос: почему не 3, ведь на конце модельного ВР мы наблюдаем еще одно сочетание лингвистических переменных BA. Не считаем такое сочетание на конце ВР за переход, т.к. перехода в

конкретную лингвистическую переменную нет. По аналогичному правилу $C(B) = 2$.

Таким образом, анализ полученных результатов верификации и валидации модели показывает, что для лингвистического временного ряда с долговременной памятью $L^* = 5$ относительная ошибка прогноза в вероятностном измерении равна 16,67%, т.е. 5 из 6 значений спрогнозированы со следующей точностью в интервале от 92% до 100% и только одно значение «угадано» как С с вероятностью 0,46%, как А с вероятностью 54%. Отдельного исследования требует изучение вопроса об уровне доверия, что планируется авторами в дальнейших исследованиях.

Первые пять значений ВР определяют долговременную память и соответственно не верифицируются. Значение итогового прогноза (12-е значение) для модельного ряда с вероятностью 100% - лингвистическая переменная А, интервальная числовая характеристика которой соответствует определенному кластеру.

Используя модельный пример, авторы показали основное преимущество инструментария ЛКА, его «работоспособность» для коротких временных рядах, что особенно важно в случаях, когда нет возможности сформировать длинные временные ряды, в том числе экономических показателей. Такие высокоточные прогнозные инструменты, как генетические алгоритмы и нейронные сети требуют и показывают хорошие результаты только на больших данных. Модель ЛКА в своей классической постановке при прогнозировании длинных временных рядов экономических показателей или природных (уровень воды в реках) от 120 и более элементов выдает относительную погрешность прогноза от 11% [4-11], что подтверждает ее работоспособность при работе с большими данными.

Отдельно отметим, что в работах [4-11] авторами предлагается использовать переход от прогнозных лингвистических переменных к

численному прогнозу, опираясь на значения исторической последовательности предшествующей прогнозному значению, но авторы считают что именно интервальный прогноз, т.е. определение принадлежности прогноза к определенному кластеру значений является достаточным, чтобы использовать инструментарий линейного клеточного автомата для прогноза так называемых фрактальных временных рядов с долговременной памятью.

Представим на основе проведенного исследования сравнительный анализ линейного клеточного автомата с клеточным автоматом в его классической постановке.

Метод CATSF заимствует идеи клеточных автоматов (КА) [15], интерпретируя временные переходы состояний аналогично взаимодействию соседних ячеек по оси времени. В классических клеточных автоматах состояние каждой ячейки представлено на сетке клеток как минимум двумерного пространства (плоскости), каждая из клеток обновляется на основе её собственного состояния и состояний соседей по заранее заданным правилам. CATSF использует похожую логику, но применяет её к временным рядам, причем:

- каждый момент времени рассматривается как ячейка в одномерном клеточном автомате;
- состояние ячейки — это кластерная метка $u_0(t)$;
- вместо пространственной сетки CATSF работает на оси времени, где порядок временных точек определяет эволюцию системы.

Таким образом, CATSF можно представить как одномерный клеточный автомат (размерностью – 1D), работающий вдоль временной оси.

В клеточных автоматах окружение ячейки состоит из соседних ячеек (например, ячейки слева и справа в 1D КА). В CATSF роль окружения

играет историческая последовательность состояний, определяемая следующим образом.

Окружение временной точки t – это последовательность:

$$S(t) = (u_0(t - L^*), \dots, u_0(t - 1))$$

где L^* (глубина памяти) – это количество прошлых состояний, влияющих на следующее предсказание.

Если $L^*=3$, то три последних состояния составляют окружение, аналогично тому, как в КА окружение может включать три соседние ячейки.

В CATSF L^* изменяется динамически в зависимости от данных, что похоже на переменный размер окружения в гетерогенных клеточных автоматах.

В классических КА обновление ячеек происходит по жёстким детерминированным правилам (например, "если три соседние ячейки имеют определённые состояния, следующая ячейка принимает определённое состояние"). В CATSF вместо этого используются вероятностные правила обновления, основанные на данных временного ряда.

Вес перехода вычисляется по формуле:

$$W(s_i(x; t)) = \frac{c(S(t)[i:] \oplus x)}{c(S(t)[i:])},$$

– это вероятность появления состояния x после определённой последовательности прошлых состояний, причем, чем чаще последовательность встречается в данных, тем выше вероятность её продолжения.

– итоговая вероятность перехода $P(x | S(t))$ получается, как сумма весов для всех возможных подпоследовательностей.

– это аналогично таблице переходов в КА, но в CATSF она основана не на жёстких правилах, а на статистических закономерностях.

Метод CATSF обновляет систему пошагово:

1. Определяется окружение $S(t)$ (временной интервал);
2. Применяются вероятностные переходы, рассчитывая $P(x | S(t))$;
3. Выбирается следующее состояние: $\hat{x}(t) = \arg \max P(x | S(t))$;
4. Прогноз следующего состояния: $S(t + 1) = S(t) \oplus \hat{x}(t)$.

Этот процесс похож на синхронное обновление в КА, где все ячейки меняются одновременно, но в CATSF обновления идут последовательно по временной оси.

Таким образом, метод CATSF связан с клеточными автоматами следующими ключевыми аналогиями:

- Временные точки $\rightarrow$ ячейки в 1D КА;
- Отрезок временного ряда $\rightarrow$ окружение в КА;
- Вероятностные переходы $\rightarrow$ правила обновления в КА;
- Долговременная память L^* $\rightarrow$ изменяющийся размер окружения в КА;
- Прогнозное состояние $\rightarrow$ эволюция КА во времени.

Таким образом, CATSF можно рассматривать как вероятностный вариант клеточного автомата, который оперирует на временной оси, а не в пространстве.

Заключение

Таким образом, можно сделать вывод, что использование методов машинного обучения как встроенного инструмента автоматизации обучения ЛКА с точки зрения комплексного подхода к моделированию прогнозного инструмента, включая разработку и использование методов K-means для обоснования количества (кластеров) цветов ЛКА обладает научной новизной. Во всех предшествующих исследовательских работах [4-11], включая разработанные программные продукты [16,17] авторы использовали трех-четырех цветный ЛКА, с привлечением экспертного подхода изложенного в первой работе В.А. Перепелицы и М.Д. Касевой.

Все три поставленные авторами цели исследования достигнуты, включая сравнительный анализ и обоснование правомерности использования для названия модели термина «линейный клеточный автомат», т.к. в публикации, в которой авторами впервые представлена прогнозная модель ЛКА [3], как и во всех последующих работах нет пояснения почему для модели использовано название «Линейный клеточный автомат». Проведя детальный анализ, в настоящем исследовании восполнен образовавшийся пробел. Подробное аналитическое описание математической модели ЛКА позволит всем читателям, заинтересовавшимся данным инструментарием, провести аналогичное самостоятельное исследование, а такие встроенные методы машинного обучения как дополненный метод локтя и силуэтный анализ расширили возможности использования ЛКА как прогнозного инструментария.

Предложенная в работе математическая модель прогноза на базе комплекса методов k-means и клеточных автоматов реализована авторами в виде программного продукта [18]. Результативность применения и использования которого планируется апробировать на временных рядах, обладающих фрактальными свойствами [19] и представить в будущих публикациях.

СПИСОК ИСТОЧНИКОВ

1. Кричевский М. Л. Интеллектуальный анализ данных в менеджменте. СПбГУАП. СПб., 2005. 208с.
2. Петерс Э. Хаос и порядок на рынках капитала. Новый аналитический взгляд на циклы, цены и изменчивость рынка. Москва: Мир., 2000. 333с.
3. 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.
4. Popova, E., de Sousa Costa, L., Kumratova, A. (2020). Hybrid Instrumental Means of Predictive Analysis of the Dynamics of Natural and Economic Processes. In: Madureira, A., Abraham, A., Gandhi, N., Varela, M. (eds) Hybrid Intelligent Systems. HIS 2018. Advances

in Intelligent Systems and Computing, vol 923. Springer, Cham. 2020. https://doi.org/10.1007/978-3-030-14347-3_4.

5. Кумратова, А. М. Математические методы и инструментальные средства прогнозирования сложных процессов / А. М. Кумратова, И. В. Розентул, А. И. Василенко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2024. – № 204. – С. 152-166. – DOI 10.21515/1990-4665-204-017. – EDN FQIPUW.

6. Tourist and recreational complex elements dynamics research data analysis qualitative methods / A. M. Kumratova, E. V. Popova, L. O. Velikanova [et al.] // IOP Conference Series: Earth and Environmental Science : 3, Volgograd, Virtual, 10–11 декабря 2020 года. – Volgograd, Virtual, 2021. – P. 012010. – DOI 10.1088/1755-1315/786/1/012010. – EDN MNDCQZ.

7. Попова Е. В. Теория нечетких множеств и клеточных автоматов как инструментарий прогноза 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.

8. Перепелица, В.А., Тамбиева Д.А. Системы с иерархической структурой управления: разработка экономико-математических и инструментальных методов / В. А. Перепелица, Д. А. Тамбиева. - Москва : Финансы и статистика, 2009. - 269 с. : ил., табл.; 20 см.; ISBN 978-5-279-03415-4

9. Замотайлова, Д. А. Применение тренд-сезонных моделей для исследования и прогнозирования временных рядов / Д. А. Замотайлова, Д. Н. Савинская, Т. А. Недогонова // Современная экономика: проблемы и решения. – 2019. – № 12(120). – С. 40-48. – DOI 10.17308/meps.2019.12/2197. – EDN ECQPYW.

10. 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.

11. Кумратова, А. М. Экономико-математическое моделирование риска в задачах управления ресурсами здравоохранения / А.М. Кумратова, Е.В. Попова, А.З. Биджиев; Кубанский государственный аграрный университет. – Краснодар: Кубанский государственный аграрный университет имени И.Т. Трубилина, 2014. – 168 с. – ISBN 978-5-94672-792-1. – EDN SMTWWB.

12. Nasir Saeed, Haewoon Nam, Mian Imtiaz Ul Haq, and Dost Muhammad Saqib Bhatti. A Survey on Multidimensional Scaling. //ACM Comput. Surv. – 2018. 51, 3, Article 47, 25 pages. <https://doi.org/10.1145/3178155>.

13. Rousseeuw P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis // Journal of Computational and Applied Mathematics. – 1987. – Т. 20. – С. 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).

14. Onumanyi AJ, Molokomme DN, Isaac SJ, Abu-Mahfouz AM. AutoElbow: An Automatic Elbow Detection Method for Estimating the Number of Clusters in a Dataset. //Applied Sciences. – 2022. № 12(15):7515. <https://doi.org/10.3390/app12157515>.

15. Wolfram, S. Statistical mechanics of cellular automata. In Reviews of Modern Physics //American Physical Society (APS). – 1983. Vol. 55, Issue 3, pp. 601–644. <https://doi.org/10.1103/revmodphys.55.601>.

16. Комплексный анализ сложных систем. Кумратова А.М., Василенко А.И., Егорчев Ф.А., Параскевов А.В. Свидетельство о регистрации программы для ЭВМ RU 2024689000, 03.12.2024. Заявка № 2024687639 от 14.11.2024

17. Программа для проведения предпрогнозного анализа временных рядов урожайности зерновых культур. Чупин Р.И., Алещенко В.В., Кумратова А.М., Попова М.И. Свидетельство о регистрации программы для ЭВМ RU 2022680968, 08.11.2022. Заявка № 2022680390 от 28.10.2022.

18. Информационная система прогноза динамики временных рядов на базе клеточного автомата с автоматизацией обучения CAFTS. Попова Е.В., Хаммуд А., Попова М.И. Свидетельство о регистрации программы для ЭВМ RU 2025611854, 23.01.2025 Заявка № 2025610320 от 10.01.2025.

19. 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.

References

1. Krichevskij M. L. Intellectual'ny'j analiz danny'x v menedzhmente. SPbGUAP. SPb., 2005. 208s.

2. Peters E'. Хаос i poriadok na ry'nkax kapitala. Novy'j analiticheskij vzglyad na cikly', ceny' i izmenchivost' ry'nka. Moskva: Mir., 2000. 333s.

3. Perepelicza, V.A., Kasaeva, M.D. Prognozirovanie prirodnogo vremennogo ryada na baze modeli kletchnogo avtomata / V.A Perepelicza, M.D. Kasaeva //Sovremennyy'e aspekty' e'konomiki. Sankt-Peterburg. – 2002. № 9. – S. 209-212.

4. Popova, E., de Sousa Costa, L., Kumratova, A. (2020). Hybrid Instrumental Means of Predictive Analysis of the Dynamics of Natural and Economic Processes. In: Madureira, A., Abraham, A., Gandhi, N., Varela, M. (eds) Hybrid Intelligent Systems. HIS 2018. Advances in Intelligent Systems and Computing, vol 923. Springer, Cham. 2020. https://doi.org/10.1007/978-3-030-14347-3_4.

5. Kumratova, A. M. Matematicheskie metody' i instrumental'ny'e sredstva prognozirovaniya slozhny'x processov / A. M. Kumratova, I. V. Rozentul, A. I. Vasilenko // Politematicheskij setevoy e'lektronny'j nauchny'j zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. – 2024. – № 204. – S. 152-166. – DOI 10.21515/1990-4665-204-017. – EDN FQIPUW.

6. Tourist and recreational complex elements dynamics research data analysis qualitative methods / A. M. Kumratova, E. V. Popova, L. O. Velikanova [et al.] // IOP Conference Series: Earth and Environmental Science : 3, Volgograd, Virtual, 10–11 dekabrya 2020 goda. – Volgograd, Virtual, 2021. – P. 012010. – DOI 10.1088/1755-1315/786/1/012010. – EDN MNDCQZ.

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

8. Perepelicza, V.A., Tambieva D.A. Sistemy' s ierarxicheskoy strukturoj upravleniya: razrabotka e'konomiko-matematicheskix i instrumental'ny'x metodov / V. A. Perepelicza, D. A. Tambieva. - Moskva : Finansy' i statistika, 2009. - 269 s. : il., tabl.; 20 sm.; ISBN 978-5-279-03415-4

9. Zamotajlova, D. A. Primenenie trend-sezonny'x modelej dlya issledovaniya i prognozirovaniya vremenny'x ryadov / D. A. Zamotajlova, D. N. Savinskaya, T. A.

Nedogonova // *Sovremennaya e'konomika: problemy' i resheniya*. – 2019. – № 12(120). – S. 40-48. – DOI 10.17308/meps.2019.12/2197. – EDN ECQPYW.

10. 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.

11. Kumratova, A. M. *E'konomiko-matematicheskoe modelirovanie riska v zadachax upravleniya resursami zdavooxraneniya* / A.M. Kumratova, E.V. Popova, A.Z. Bidzhiev; Kubanskij gosudarstvenny'j agrarny'j universitet. – Krasnodar: Kubanskij gosudarstvenny'j agrarny'j universitet imeni I.T. Trubilina, 2014. – 168 s. – ISBN 978-5-94672-792-1. – EDN SMTWWB.

12. Nasir Saeed, Haewoon Nam, Mian Imtiaz Ul Haq, and Dost Muhammad Saqib Bhatti. A Survey on Multidimensional Scaling. // *ACM Comput. Surv.* – 2018. 51, 3, Article 47, 25 pages. <https://doi.org/10.1145/3178155>.

13. Rousseeuw P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis // *Journal of Computational and Applied Mathematics*. – 1987. – T. 20. – C. 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).

14. Onumanyi AJ, Molokomme DN, Isaac SJ, Abu-Mahfouz AM. AutoElbow: An Automatic Elbow Detection Method for Estimating the Number of Clusters in a Dataset. // *Applied Sciences*. – 2022. № 12(15):7515. <https://doi.org/10.3390/app12157515>.

15. Wolfram, S. Statistical mechanics of cellular automata. In *Reviews of Modern Physics* // *American Physical Society (APS)*. – 1983. Vol. 55, Issue 3, pp. 601–644. <https://doi.org/10.1103/revmodphys.55.601>.

16. Kompleksny'j analiz slozhny'x sistem. Kumratova A.M., Vasilenko A.I., Egorchev F.A., Paraskevov A.V. Svidetel'stvo o registracii programmy' dlya E'VM RU 2024689000, 03.12.2024. Zayavka № 2024687639 ot 14.11.2024

17. Programma dlya provedeniya predprognoznogo analiza vremenny'x ryadov urozhajnosti zernovy'x kul'tur. Chupin R.I., Aleshhenko V.V., Kumratova A.M., Popova M.I. Svidetel'stvo o registracii programmy' dlya E'VM RU 2022680968, 08.11.2022. Zayavka № 2022680390 ot 28.10.2022.

18. Informacionnaya sistema prognoza dinamiki vremenny'x ryadov na baze kletochnogo avtomata s avtomatizaciej obucheniya SAFTS. Popova E.V., Xammud A., Popova M.I. Svidetel'stvo o registracii programmy' dlya E'VM RU 2025611854, 23.01.2025. Zayavka № 2025610320 ot 10.01.2025.

19. 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.