

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ: СОВРЕМЕННЫЕ ПОДХОДЫ И ПЕРСПЕКТИВЫ

Захарова О.И., Бедняк С.Г., Квачахия И.З.

Поволжский государственный университет телекоммуникаций и информатики, Самара, РФ

E-mail: o.zaharova@psuti.ru

Статья рассматривает использование методов машинного обучения для прогнозирования временных рядов, что является важной задачей в различных областях, таких как экономика, финансы, метеорология и бизнес-аналитика. Традиционные статистические методы, включая модели ARIMA, показали ограниченную применимость при работе с нелинейными временными данными. В связи с этим современные подходы, основанные на машинном обучении, приобрели большое значение. Основное внимание в статье уделено таким методам, как искусственные нейронные сети, регрессия ближайших соседей и машины опорных векторов, которые демонстрируют высокую точность и устойчивость при работе с большими объемами данных. Цель исследования заключается в анализе существующих подходов к прогнозированию временных рядов с использованием машинного обучения, выявлении преимуществ и недостатков различных алгоритмов, а также в обсуждении методов для одношагового и многошагового прогнозирования. Статья подчеркивает значимость локального обучения как гибкого инструмента для работы с нестационарными данными и предлагает перспективы для дальнейшего развития в этой области.

Ключевые слова: прогнозирование временных рядов, машинное обучение, ARIMA модели, статистические методы, стратегии прогнозирования

Введение

Временной ряд представляет собой последовательность значений наблюдаемой переменной, зарегистрированных через одинаковые интервалы времени. Исследование таких рядов проводится с целью прогнозирования будущих значений на основе прежних наблюдений, понимания основных закономерностей или краткого описания ключевых характеристик ряда. Данная статья фокусируется на проблеме прогнозирования, которое играет решающую роль во многих сферах, включая экономику, финансы, бизнес-аналитику, метеорологию и телекоммуникации. Важным аспектом в прогнозировании является длина горизонта прогнозирования, так как многоэтапное прогнозирование усложняется накоплением ошибок, снижением точности и ростом неопределенности.

Традиционно в прогнозировании применялись линейные статистические методы, такие как модели ARIMA (Autoregressive Integrated Moving Average), однако с конца 1970-х годов стало очевидно, что данные методы не всегда подходят для практических применений [1; 2]. В это время были предложены нелинейные модели, такие как билинейная модель, пороговая авторегрессионная модель и модель авторегрессионной условной гетероскедастичности (Autoregressive Conditional Heteroscedasticity, ARCH) [3]. Однако по сравнению с линейными моделями, исследование и применение нелинейных моделей находятся на начальном этапе [4].

На протяжении последних двух десятилетий модели машинного обучения стали набирать популярность и зарекомендовали себя как конкурентоспособные альтернативы классическим статистическим моделям [5]. Эти модели, называемые также «черными ящиками» или моделями, управляемыми данными, являются примерами непараметрических нелинейных моделей, использующих исключительно исторические данные для выявления взаимосвязей между прошлыми и будущими значениями [6; 7]. Например, исследования П.Ж. Вербоса показали, что искусственные нейронные сети (ИНС) превосходят традиционные подходы, такие как линейные регрессии и модели Бокса-Дженкинса. А. Лапедес и Р. Фарбер также обнаружили, что ИНС могут быть успешно использованы для моделирования и прогнозирования нелинейных временных рядов [8]. В дальнейшем появились и другие методы, такие как деревья решений, машины опорных векторов и регрессия ближайших соседей.

Настоящая работа представляет обзор методов машинного обучения с акцентом на формализацию задач одношагового прогнозирования и применение методов локального обучения, фрактальных методов и обучения с учителем. В работе систематизированы современные подходы к переходу от одношагового к многошаговому прогнозированию и исследованы возможности их применения с использованием методов локального обучения.

Целью данной работы является изучение и систематизация современных подходов машинного обучения для прогнозирования временных рядов, с акцентом на анализ методов локального обучения, их возможностей и ограничений, а также разработка рекомендаций по применению данных методов в задачах прогнозирования.

Структура материала состоит из трех основных разделов. В разделе «Прогнозирование и моделирование» рассмотрены основные концепции моделирования временных рядов и формализована задача прогнозирования как задача «вход-выход». Раздел «Подходы машинного обучения для моделирования временных зависимостей», посвящен обсуждению роли машинного обучения в создании точных предсказателей на основе наблюдаемых данных и знакомит с понятием локального обучения.

Прогнозирование и моделирование

Существуют две главные интерпретации задачи прогнозирования на основе исторических данных [9; 10]. Согласно теории статистического прогнозирования, наблюдаемая последовательность является одной из возможных реализаций случайного процесса, где случайность возникает из-за множества независимых степеней свободы, взаимодействующих линейно [5]. В то же время, теория динамических систем предлагает иную перспективу, предполагая, что кажущееся случайным поведение может быть результатом функционирования детерминированных систем с малым количеством степеней свободы, взаимодействующих нелинейно [8]. Такое сложное и нерегулярное поведение называется детерминированным хаосом.

Многие типы временных рядов, получаемых в экспериментальных данных, могут быть изучены с помощью теории динамических систем. В данном подходе временной ряд рассматривается как динамическая система, в которой состояние $s(t)$ изменяется во времени в пространстве состояний $\Gamma \subset R^n$ по определенному закону:

$$s(t) = F^t(s(0)), \quad (1)$$

где $F: \Gamma \rightarrow \Gamma$ – карта, представляющая динамику;

F^t – итерационное применение отображения F , то есть t -кратное последовательное применение F :

$$F^t(s(0)) = F(F(\dots F(s(0))\dots)),$$

где функция F применяется t раз;

$s(0) \in \Gamma$ – начальное состояние системы;

$s(t) \in \Gamma$ – состояние системы в момент времени t ;
 t – целочисленный дискретный индекс времени.

Эта формула описывает эволюцию состояния системы во времени, где поведение системы полностью определяется динамическим правилом F и начальным состоянием $s(0)$.

В контексте анализа временных рядов наблюдаемое значение в момент времени t , как правило, не совпадает напрямую с внутренним состоянием $s(t)$. Вместо этого наблюдаемое значение определяется функцией измерения G , которая связывает состояние системы с наблюдаемыми данными:

$$y_t = G(s(t)), \quad (2)$$

где $G: \Gamma \rightarrow RD$ – функция измерения, которая отображает внутреннее состояние системы $s(t)$ в наблюдаемую величину y_t ;

D – размерность выходной наблюдаемой серии.

В данной статье рассматривается случай $D = 1$, то есть одномерный временной ряд.

Таким образом, функция G задает связь между состоянием системы $s(t)$ и наблюдаемой величиной y_t . Если G неизвестна, восстановление динамики системы на основании наблюдаемых данных становится нетривиальной задачей.

Формула (1) описывает, как состояние системы изменяется во времени в пространстве состояний Γ .

Формула (2) связывает внутреннее состояние системы $s(t)$ с наблюдаемым значением y_t через функцию измерения G .

Функции F и G неизвестны, поэтому в общем случае нельзя надеяться на восстановление состояния в его первоначальном виде. Однако имеется возможность воссоздать пространство состояний, которое в некотором смысле эквивалентно исходному.

Когда исследуют динамические системы, часто оказывается так, что важная информация о системе скрыта в наблюдаемых данных. Это особенно актуально, если непосредственный доступ к внутренним переменным системы отсутствует. Задача реконструкции пространства состояний решает проблему восстановления динамики неопределенной системы через те же наблюдаемые данные. *Задача реконструкции пространства состояний* состоит в восстановлении состояния, когда единственная доступная информация содержится в последовательности наблюдений y_t . Основная идея заключается в использовании метода вложений (Embedding), предложенного Ф. Такенсом (Takens' Theorem). Этот метод позволяет, используя лишь один временной ряд наблюдений, воссоздать фазовое пространство, эквивалентное исходному, следуя ключевым шагам: (*карта реконструкции*) $\Phi(s(t))$:

$$\Phi(s(t)) = \{G(F^{-d}(s(t))), \dots, G(F^{-1}(s(t))), G(s(t))\}.$$

Это выражение можно записать через наблюдаемые значения временного ряда:

$$\Phi(s(t)) = \{y_t - d, \dots, y_t - 1, y_t\},$$

где $\Phi: \Gamma \rightarrow R^m$ – карта реконструкции, отображающая состояния системы в реконструированное фазовое пространство размерности m ;

G – функция измерения, связанная с наблюдаемыми данными y ,

F^{-k} – обратное (временное) применение карты F , соответствующее состоянию системы на k -шагов назад;

d – временная задержка, определяющая шаг между элементами реконструированного вектора;

m – размерность реконструированного пространства (число используемых значений), называемая вложением или порядком. Согласно теореме Такенса, $m > 2d$, что гарантирует однозначное восстановление динамики системы;

t – индекс временного ряда, представляющий номер отсчета в дискретной последовательности $\{y_t\}$.

Основным следствием применения карты реконструкции Φ является то, что если Φ является вложением, то в пространстве восстановленных векторов индуцируется гладкая динамика f , задаваемая следующим уравнением:

$$y_t = f(y_{t-d}, y_{t-d-1}, \dots, y_{t-d-n+1}). \quad (3)$$

где $f: R^n \rightarrow R$ – отображение, описывающее динамическую зависимость текущего наблюдения y_t от n -мерного вектора запаздывающих значений;

n – размерность реконструированного пространства, которая определяется количеством включенных задержек и должна быть достаточно большой ($n > 2dn$) для корректной реконструкции динамики;

d – шаг задержки (временной интервал между значениями в векторе). Это означает, что реконструированные состояния в пространстве R^n могут быть использованы для оценки функции f , которая описывает динамическое поведение временного ряда. Таким образом, f может выступать альтернативой исходным отображениям F и G для решения различных задач, включая качественное описание динамики, прогнозирование временных рядов и анализ их свойств.

Представление (3) не учитывает никакой шумовой компонент, поскольку оно предполагает, что детерминированный процесс f может точно описать временные ряды. Однако это просто один

из возможных способов представления явления временного ряда, и любое альтернативное представление не должно быть отвергнуто априори. На самом деле, если предположить, что доступ к точной модели функции f нет, то вполне разумным решением будет расширить детерминистическую формулировку (3) до статистической нелинейной авторегрессионной формулировки (NAR).

$$y_t = f(y_{t-d}, y_{t-d-1}, \dots, y_{t-d-n+1}) + w(t), \quad (4)$$

где недостающая информация объединяется в шумовой член. В остальной части статьи мы будем ссылаться на формулировку (4) как на общее представление временного ряда, которое включает в себя в качестве частного случая также случай (3).

Успех подхода к реконструкции на основе набора наблюдаемых данных зависит от выбора гипотезы, которая аппроксимирует f , выбора порядка n и времени запаздывания d .

В следующем разделе будет рассмотрена проблема моделирования f , если предполагать, что значения n и d доступны априори.

Подходы машинного обучения для моделирования временных зависимостей

Обучение с учителем. Формулировка вложения в (4) указывает на то, что при наличии исторической последовательности S задача одношагового предсказания может быть решена как задача обучения с учителем. Обучение с учителем включает в себя создание модели, с опорой на множество наблюдений, которая устанавливает связь между входными переменными и одной или несколькими зависимыми выходными переменными. После того, как модель по отображению (4) построена, ее можно применять для одношагового предсказания. В этом случае используются n предшествующих значений ряда, и задача предсказания представляется как обобщенная регрессионная задача, что показано на рисунке 1.

Общий подход к моделированию явления «вход/выход» со скалярным выходом и векторным входом основывается на наличии набора наблюдаемых пар, обычно называемых обучающим множеством.

При прогнозировании обучающее множество формируется на основании исторического ряда S путем создания матрицы $[(N - n - 1) \times n]$ входных данных:

$$X = \begin{bmatrix} y_{N-1} & \cdots & y_{N-n} \\ \vdots & \ddots & \vdots \\ y_n & \cdots & y_1 \end{bmatrix},$$

и выходного вектора $[(N - n - 1) \times 1]$

$$Y = \begin{bmatrix} y_N \\ \dots \\ y_{N+1} \end{bmatrix}.$$

Для простоты мы предполагаем здесь, что временная задержка $d = 0$, то есть значения векторов в матрице X берутся подряд без пропусков. В дальнейшем будем ссылаться на i -ый ряд матрицы X , который представляет собой запаздывающий вектор временного ряда – вектор состояния, описывающий n последовательных значений временного ряда. Этот вектор отражает состояние временного ряда в момент времени $t - i + 1$.

Инстанцирование с локальным обучением. Прогнозирование на шаг вперед заключается в предсказании значения выхода при заданном подмножестве прошлых наблюдаемых значений (также обозначаемом как запрос). Машинное обучение обеспечивает теоретическую основу для оценки по наблюдаемым данным подходящей модели временной зависимости f . Из-за невозможности рассмотреть здесь весь современный уровень машинного обучения в прогнозировании временных рядов, далее мы более конкретно рассмотрим методы локального обучения. Этот выбор обусловлен следующими причинами:

1. Гибкость подхода. Локальные методы обучения не требуют строгих предположений о характере распределения данных или специфике процесса, лежащего в основе временного ряда. Это делает их пригодными для работы с реальными данными, которые часто имеют высокую степень нестационарности и шумности.

2. Отсутствие необходимости в глобальной модели. Методы локального обучения не зависят от построения общей модели для описания всех данных. Вместо этого они используют только ближайшие данные для прогноза, что снижает риск переобучения на данных, содержащих выбросы или шум.

3. Устойчивость к нестационарности. Локальное обучение эффективно работает с временными рядами, характеристики которых изменяются со временем. Методы легко адаптируются к новым условиям, что делает их особенно полезными для динамических и изменяющихся систем.

4. Интерпретируемость. Результаты локальных методов, таких как метод ближайших соседей, могут быть легко интерпретированы, поскольку они основываются на сравнении с известными ранее наблюдениями.

5. Возможность онлайн-обучения. Локальные методы обучения подходят для сценариев,

когда данные поступают в режиме реального времени. Такие методы позволяют обновлять обучающую выборку без необходимости полного переобучения модели, что снижает вычислительные затраты.

Локальное обучение не требует предварительных знаний о процессе, который лежит в основе данных. Например, оно не делает предположений о наличии общей функции, описывающей данные, и не имеет никакой связи со свойствами шума. Единственная доступная информация – это набор входных и выходных наблюдений. Эта особенность, главным образом важна для реальных данных, где проблемы отсутствующих характеристик, нестационарности и ошибок измерения делают подход, основанный на данных и свободный от предположений, особенно привлекательным.

Метод локального обучения легко справляется с задачами обучения в режиме онлайн, когда количество обучающих данных растет с течением времени. В таком случае локальное обучение просто добавляет новые точки в набор данных и не требует трудоемкого переобучения при появлении новых данных.

Метод локального обучения может работать с изменяющимися во времени конфигурациями, когда случайный процесс, лежащий в основе данных, является нестационарным. В таких случаях соседей целесообразно определять не только по пространству, но и по времени. Для каждой новой точки запроса соседями становятся не только те данные, которые имеют схожие входные параметры, но и те, которые были собраны недавно. Таким образом, временная переменная становится еще одной важной характеристикой, которую необходимо учитывать для точного прогнозирования.

В продолжение мы рассмотрим пример методов локального обучения: метод ближайших соседей и фрактальные методы анализа временных рядов.

Фрактальные методы анализа временных рядов. Помимо методов локального обучения, фрактальные методы анализа временных рядов предоставляют дополнительные возможности для изучения сложных динамических систем. Эти подходы фокусируются на выявлении самоподобия и масштабных закономерностей, что делает их особенно полезными в задачах прогнозирования и анализа временных данных.

Фрактальные методы анализа временных рядов являются мощным инструментом для исследования сложных динамических систем. Они позволяют выявлять скрытые структуры и закономерности в данных, которые часто не удается

обнаружить с помощью традиционных методов. Фрактальные подходы основываются на концепции самоподобия, характеризующей временные ряды, и широко применяются в областях, таких как экономика, биология, геофизика и телекоммуникации.

Основные принципы фрактального анализа:

1. Самоподобие временного ряда. Временной ряд считается фрактальным, если его структура остается неизменной при изменении масштаба времени. Это свойство описывается фрактальной размерностью D , которая может быть рассчитана через различные методы, такие как метод покрытия, метод ячеек или метод подсчета боксов (box-counting).

2. Гистограмма распределения шкал (Hurst Exponent). Экспонента Херста H является важной характеристикой фрактальных временных рядов. Она показывает степень долгосрочной памяти временного ряда:

$$\frac{R}{S} = (NH),$$

где R – диапазон изменений временного ряда;

S – стандартное отклонение;

N – длина временного окна;

$H \in [0,1]$ – экспонента Херста.

3. Фрактальная размерность по методу Корнфелда (также известная как корреляционная размерность). Для вычисления фрактальной размерности D_2 используется корреляционная сумма:

$$C(r) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^{Nn} \Theta(r - \|x_i - x_j\|),$$

где $C(r)$ – корреляционная сумма;

r – радиус окрестности;

Θ – функция Хевисайда;

$\|x_i - x_j\|$ – расстояние между точками x_i и x_j . Размерность D_2 оценивается через наклон логарифмического графика зависимости $C(r)$ от r :

$$D_2 = \lim_{r \rightarrow 0} \frac{\log C(r)}{\log r}.$$

4. Детрендрованный флуктуационный анализ (DFA). DFA позволяет оценить сложность временного ряда, исключив влияние трендов. Основные этапы:

1) Разделить временной ряд на сегменты длиной n , для каждого сегмента вычислить локальный тренд методом наименьших квадратов, рассчитать флуктуации $F(n)$:

$$F(n) = \sqrt{\frac{1}{N} \sum_{i=1}^N [x(i) - T(i)]^2},$$

где $T(i)$ – локальный тренд, построить зависи-

мость $F(n)$ от n на логарифмическом графике.

Фрактальные методы позволяют:

– диагностировать наличие долгосрочных зависимостей в данных;

– выявлять масштабные структурные закономерности, что особенно полезно для сложных систем с многомерной динамикой;

– анализировать свойства нестационарных временных рядов, часто встречающихся в финансовых и телекоммуникационных данных.

Ближайший сосед. Метод ближайшего соседа (Nearest Neighbors Algorithm, NN) – это самый простой пример локальной аппроксимации, который используется для прогнозирования временных рядов [11]. Суть метода в том, чтобы найти в наборе данных точку, ближайшую к текущему состоянию, и предположить, что текущее состояние будет развиваться аналогично найденному соседу.

На рисунке 1 представлен пример одношагового прогнозирования по ближайшим соседям. Предположим, что у нас имеется временной ряд y_t до момента времени $t-1$, и мы хотим предсказать следующее значение ряда. После выбора определенной размерности n , например, $n=6$, подход ближайшего соседа ищет шаблон в прошлом, который является наиболее похожим, в данной метрике, на модель $\{y_{t-6}, y_{t-5}, \dots, y_{t-1}\}$ (пунктирная линия). Если ближайшим шаблоном является, например, $\{y_{t-16}, y_{t-15}, \dots, y_{t-11}\}$, то прогнозы $\hat{y}_t$, возвращаемые методом NN, – это значение y_{t-10} (черная точка).

Необходимо предсказать завтрашнюю погоду в Москве, имея временной ряд наблюдений. Предположим, размерность $n=1$ (одномерный шаблон). Метод NN включает следующие шаги:

1. Найти в исторических данных шаблон погоды, наиболее похожий на текущие условия. Например, это может быть погода 10 сентября 2020 года.

2. Выбрать из данных погоду на следующий день после найденного шаблона (например, случайно дождливый день).

3. Прогнозировать завтрашнюю погоду как значение, соответствующее следующему дню в историческом шаблоне.

Метод можно расширить, учитывая большее количество ближайших соседей или используя аппроксимацию более высокого порядка.

Таким образом, методы машинного обучения предоставляют мощные инструменты для решения задачи прогнозирования временных рядов, обеспечивая значительные преимущества по сравнению с традиционными статистическими подходами. В частности, методы локального об-

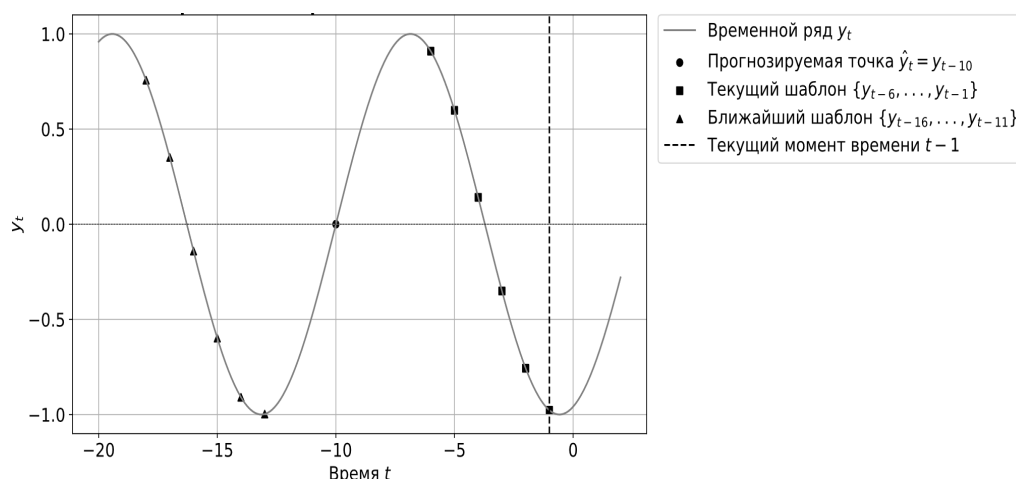


Рисунок 1. Пример одношагового прогнозирования методом ближайших соседей

учения, такие как ближайшие соседи и локальная регрессия, показывают свою эффективность в работе с реальными данными, где использование традиционных моделей может быть связано с определенными ограничениями.

Заключение

В статье приведен подробный анализ применения методов машинного обучения для предсказания временных рядов, с выделением их основных преимуществ по сравнению с традиционными статистическими моделями. Установлено, что современные алгоритмы, включая искусственные нейронные сети, машины опорных векторов и регрессию ближайших соседей, демонстрируют более точные результаты при работе с нелинейными и сложными данными.

В статье проанализированы преимущества и ограничения как традиционных линейных методов (например, ARIMA), так и современных моделей машинного обучения. Показано, что последние, включая искусственные нейронные сети, обладают способностью учитывать сложные нелинейные зависимости, которые трудно моделировать с помощью классических методов.

Кроме того, выделены методы локального обучения, такие как метод ближайших соседей, фрактальные методы анализа временных рядов, а также описаны их преимущества: гибкость, способность работать с нестационарными данными и поддержка онлайн-обучения.

Рассмотрены основные трудности многошагового прогнозирования, такие как накопление ошибок и снижение точности при увеличении горизонта прогноза.

Важно подчеркнуть, что несмотря на заметное повышение степени точности предсказаний при использовании методов машинного обучения,

выбор подходящего алгоритма и его параметров требует тщательной настройки в зависимости от специфики данных. Перспективные исследования должны быть направлены на оптимизацию существующих моделей, создание гибридных решений и дальнейшего распространения применения машинного обучения в анализе больших данных и для решения задач реального времени.

Литература

1. Hamilton J.D. Time Series Analysis. Princeton: Princeton University Press, 1994. 813 p.
2. Захарова О.И., Квачахия И.З. Разработка программных интеллектуальных средств для анализа и прогнозирования рынка ценных бумаг // Вестник науки. 2024. Т. 3, № 6 (75). С. 1367–1371.
3. Engle R.F. GARCH 101: the use of ARCH/GARCH models in applied econometrics // Journal of Economic Perspectives. 2001. Vol. 15, no. 4. P. 157–168.
4. Андерсон Т.В. Статистический анализ временных рядов. М.: Мир, 1976. 756 с.
5. Box G.E.P., Jenkins G.M., Reinsel G.C. Time Series Analysis: Forecasting and Control. 4th Ed. Hoboken: John Wiley & Sons, 2009. 303 p.
6. Zhang G., Qi M. Neural network forecasting for seasonal and trend time series // European Journal of Operational Research. 2005. Vol. 160, no. 2. P. 501–514.
7. Shumway R.H., Stoffer D.S. Time Series Analysis and its Applications with R Examples. Springer Science+Business Media, 2011. 202 p. DOI: 10.1007/978-1-4419-7865-3
8. Takens F. Detecting strange attractors in turbulence // Lecture Notes in Mathematics. 1981. Vol. 898. P. 366–381. DOI: 10.1007/BFb0091924
9. Privalsky V. Practical Time Series Analysis in Nat-

- ural Sciences. Springer, 2023. DOI: 10.1007/978-3-031-16891-8
10. Geurts M., Box G.E.P., Jenkins G.M. Time series analysis: forecasting and control // Journal of Marketing Research. 1977. Vol. 14, no. 2. P. 269. DOI: 10.2307/3150485
11. Категориальный анализ и оценка поведения клиентов для прогнозирования рыночных от-

ношений / М.А. Кораблин [и др.] // Информационные технологии. 2009. № 11. С. 47–50.

Дата поступления 12.09.2024

Дата принятия 14.11.2024

Захарова Оксана Игоревна, к.т.н., доцент, доцент кафедры информационных систем и технологий (ИСТ) Поволжского государственного университета телекоммуникаций и информатики (ПГУТИ). 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 906 343-25-21. E-mail: o.zaharova@psuti.ru

Бедняк Светлана Геннадьевна, к.п.н., доцент, доцент кафедры ИСТ ПГУТИ. 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 903 308-29-88. E-mail: s.bednyak@psuti.ru

Квачахия Ираклий Зазович, магистрант кафедры ИСТ ПГУТИ. 443090, Российская Федерация, г. Самара, Московское шоссе, 77. Тел. +7 987 949-50-32. E-mail: irakli.kvachakhia2000@gmail.com

UDC 681.518.25

DOI: 10.18469/ikt.2024.22.3.08

USING MACHINE LEARNING METHODS FOR TIME SERIES FORECASTING: CURRENT APPROACHES AND PERSPECTIVES

Zakharova O.I., Bednyak S.G., Kvachakhia I.Z.

Povolzhskiy State University of Telecommunications and Informatics, Samara, Russian Federation

E-mail: o.zaharova@psuti.ru

The article discusses the use of machine learning techniques for time series forecasting, which is an important task in various fields of life, such as economics, finance, meteorology and business analytics. Traditional statistical methods, including ARIMA models, have demonstrated limited applicability when dealing with nonlinear temporal data. In this regard, modern approaches based on machine learning have gained an actual importance. The main focus of the article is made on such methods as artificial neural networks, nearest neighbor regression and support vector machines, which demonstrate high accuracy and stability degree when working with massive datasets. The purpose of the study is to analyze existing approaches to time series forecasting using machine learning, identify advantages and disadvantages of different algorithms, and discuss methods for single-step and multi-step forecasting. The article emphasizes importance of local learning as a flexible tool for dealing with non-stationary data and also offers prospects for further development in this area as well.

Keywords: *time series forecasting, machine learning, ARIMA models, statistical methods, forecasting strategies*

Zakharova Oksana Igorevna, Povolzhskiy State University of Telecommunications and Informatics, 77, Moscovskoe shosse, Samara, 443090, Russian Federation; Associated Professor of Information Systems and Technologies Department, PhD in Technical Science. Tel. +7 906 343-25-21. E-mail: o.zaharova@psuti.ru

Bednyak Svetlana Gennadievna, Povolzhskiy State University of Telecommunications and Informatics, 77, Moscovskoe shosse, Samara, 443090, Russian Federation; Associated Professor of Information Systems and Technologies Department, PhD in Pedagogical Sciences. Tel. +7 903 308-29-88. E-mail: s.bednyak@psuti.ru

Kvachakhia Irakli Zazovich, Povolzhskiy State University of Telecommunications and Informatics, 77, Moscovskoe shosse, Samara, 443090, Russian Federation; Master's Degree Student of Information Systems and Technologies Department. Tel. +7 987 949-50-32. E-mail: irakli.kvachahia2000@gmail.com

References

1. Hamilton J.D. *Time Series Analysis*. Princeton: Princeton University Press, 1994, 813 p.
2. Zakharova O.I., Kvachakhia I.Z. Development of program intellectual means for the analysis and forecasting of the securities market. *Vestnik nauki*, 2024, vol. 3, no. 6 (75), pp. 1367–1371. (In Russ.)
3. Engle R.F. GARCH 101: the use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives*, 2001, vol. 15, no. 4, pp. 157–168.
4. Anderson T.W. *Statistical Analysis of Time Series*. Moscow: Mir, 1976, 756 p. (In Russ.)
5. Box G.E.P., Jenkins G.M., Reinsel G.C. *Time Series Analysis: Forecasting and Control*. 4nd Ed. Hoboken: John Wiley & Sons, 2009, 303 p.
6. Zhang G., Qi M. Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, 2005, vol. 160, no. 2, pp. 501–514.
7. Shumway R.H., Stoffer D.S. *Time Series Analysis and its Applications with R Examples*. Springer Science+Business Media, 2011, 202 p. DOI: 10.1007/978-1-4419-7865-3
8. Takens F. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*, 1981, vol. 898, pp. 366–381. DOI: 10.1007/BFb0091924
9. Privalsky V. *Practical Time Series Analysis in Natural Sciences*. Springer, 2023. DOI: 10.1007/978-3-031-16891-8
10. Geurts M., Box G.E.P., Jenkins G.M. Time series analysis: forecasting and control. *Journal of Marketing Research*. 1977, vol. 14, no. 2, pp. 269. DOI: 10.2307/3150485
11. Korablin M.A. et al. Category analysis and estimation of client's behavior for prognostication of market relations. *Informacionnye tekhnologii*, 2009, no. 11, pp. 47–50. (In Russ.)

Received 12.09.2024

Accepted 14.11.2024