

РАЗРАБОТКА АВТОРЕГРЕССИОННОЙ НЕЙРОСЕТЕВОЙ МОДЕЛИ ДЛЯ ПРОГНОЗИРОВАНИЯ АВАРИЙНОСТИ НА ТЕРРИТОРИИ ХАБАРОВСКОГО КРАЯ

Е. В. Шохирев¹, П. П. Володькин², В. А. Лазарев³

Тихоокеанский государственный университет, Хабаровск, Россия

¹ e-mail: shohirev_kna@mail.ru

² e-mail: 004167@togudv.ru

³ e-mail: 000136@togudv.ru

А. В. Константинов

Институт горного дела Дальневосточного отделения Российской академии наук, Хабаровск, Россия

e-mail: alex-sdt@yandex.ru

Аннотация. Актуальность настоящего исследования обоснована тем, что проблема аварийности на автомобильных дорогах является серьезной социальной и экономической проблемой, одним из возможных направлений решения которой заключается в проведении профилактики и предупреждения дорожной аварийности. Основой по предотвращению и в последующем снижению аварийности на дорогах может послужить анализ произошедших ранее дорожно-транспортных происшествий, результаты которого впоследствии, с применением современных технологий можно использовать, как базу для моделирования случаев аварийности и возможности их недопущения в дальнейшем. Решением данной задачи может являться создание модели прогнозирования на основе статистических данных с применением машинного обучения в анализе дорожной аварийности, что будет эффективным при обработке данных и принятию обоснованных решений для повышения безопасности на дорогах и сможет значительно улучшить точность прогнозов, что впоследствии может быть направлено на повышение уровня безопасности. Целью исследования является сравнение фактических показателей аварийности в Хабаровском крае за период с 01.01.2015 по 30.11.2023 гг. с результатами модели прогнозирования аварийности путем обучения и апробации на основе рекуррентной нейронной сети с использованием методов машинного обучения.

В исследовании применяются научные методы статистического моделирования временных рядов, анализ признакового пространства данных. Для моделирования и последующего прогнозирования использовалась нейронная сеть рекуррентного типа, с применением авторегрессионного подхода.

Научная новизна представленного исследования выражена в виде разработанной авторегрессионной нейросетевой модели прогнозирования аварийности для выявления зависимостей и закономерностей данных, которая позволит улучшить качество прогнозирования аварийности, с учетом возможности обработки большого объема данных с информацией о факторах, влияющих на аварийность.

Дальнейшие исследования будут направлены на совершенствование модели прогнозирования, с применением большего объема информации в исходных данных.

Новые научные результаты выражены в виде разработанной модели прогнозных значений ДТП, которая воспроизводит общие тенденции и предсказывает пики количества ДТП за указанный период времени. Модель, обученная на данных из обучающей выборки, способна предсказывать количество ДТП на данных тестовой выборки.

Представленная в исследовании модель прогнозирования в будущем может быть адаптирована для прогнозирования количества и видов ДТП, участников и пострадавших.

Применение результатов исследования направлено на улучшение качества профилактики дорожной аварийности и предотвращения аварийных ситуаций, используя данные модели прогнозирования.

Ключевые слова: дорожно-транспортное происшествие, анализ аварийности, безопасность дорожного движения, нейронная сеть, моделирование, прогнозирование аварийности.

Для цитирования: Разработка авторегрессионной нейросетевой модели для прогнозирования аварийности на территории Хабаровского края / Е. В. Шохирев [и др.] // Интеллект. Инновации. Инвестиции. – 2025. – № 3. – С. 107–120. – <https://doi.org/10.25198/2077-7175-2025-3-107>.

Original article

DEVELOPMENT OF AN AUTOREGRESSIVE NEURAL NETWORK MODEL FOR PREDICTING ACCIDENTS IN THE Khabarovsk Territory

E. V. Shokhirev¹, P. P. Volodkin², V. A. Lazarev³

Pacific National University, Khabarovsk, Russia

¹ e-mail: shohirev_kna@mail.ru

² e-mail: 004167@togudv.ru

³ e-mail: 000136@togudv.ru

A. V. Konstantinov

Mining Institute of the Far Eastern Branch of the Russian Academy of Sciences, Khabarovsk, Russia

e-mail: alex-sdt@yandex.ru

Abstract. The relevance of this study is justified by the fact that the problem of accidents on highways is a serious social and economic problem, one of the possible solutions to which is the prevention and prevention of road accidents. The basis for the prevention and subsequent reduction of accidents on the roads can be the analysis of previous road accidents, the results of which can be used later, with the use of modern technology as a basis for modeling of accidents and the possibility of preventing them in the future. The solution to this problem can be the creation of a prediction model based on statistical data with the use of machine learning in the analysis of road accidents, which will be effective in processing data and making informed decisions to improve road safety and can significantly improve the accuracy of forecasts, which can then be aimed at improving the level of safety. The purpose of the study is to compare the actual accident rates in Khabarovsk Krai for the period from 01.01.2015 to 30.11.2023 with the results of the accident prediction model by training and validation based on recurrent neural network using machine learning methods.

The study applies scientific methods of statistical modeling of time series, analysis of the feature space of data. A neural network of recurrent type was used for modeling and subsequent forecasting, using an autoregressive approach.

Scientific novelty in the form of development of autoregressive neural network model of accident rate forecasting to identify dependencies and patterns of data, which can improve the quality of accident rate forecasting, considering also the possibility of processing a large amount of data with information about the factors affecting the accident rate.

Further research will be aimed at improving the forecasting model, using more information in the input data.

Scientific innovation of the presented research is expressed in the form of developed autoregressive neural network model of accident rate forecasting to identify dependencies and patterns of data, which will improve the quality of accident rate forecasting, taking into account the possibility of processing a large amount of data with information about the factors affecting the accident rate. The prediction model presented in the study can be adapted in the future to predict the number and types of crashes, participants and casualties.

The application of the results of the study aims to improve the quality of road accident prevention and crash avoidance using these prediction models.

Key words: road accidents, accident analysis, traffic safety, neural network, modeling, accident prediction.

Cite as: Shokhirev, E. V., Volodkin, P. P., Lazarev, V. A., Konstantinov, A. V. (2025) [Development of an autoregressive neural network model for predicting accidents in the Khabarovsk Territory]. *Intellect. Innovacii. Investicii* [Intellect. Innovations. Investments]. Vol. 3, pp. 107–120.— <https://doi.org/10.25198/2077-7175-2025-3-107>.

Введение

Современное развитие экономики невозможно без эффективной транспортной системы [10]. Автомобильный транспорт в этой системе играет ключевую роль в грузовых и пассажирских перевозках, обслуживает различные сектора экономики, удовлетворяя потребности производства и сельского хозяйства, что увеличивает скорость и объемы товарооборота, а также повышает мобильность населения и укрепляет

экономические и культурные связи.

Аварийность на дорогах является серьезной социальной и экономической проблемой [9]. Дорожно-транспортные происшествия (далее – ДТП) приводят к травмам и смертям, а также наносят значительный материальный ущерб [3]. Это негативно сказывается на эффективности транспортных услуг.

В настоящее время на территории Российской Федерации действует федеральный проект «Безопас-

ность дорожного движения»¹. Несмотря на предпринимаемые меры и тенденции к снижению аварийности и ее последствий на дорогах, согласно данным, опубликованным на официальном сайте МВД², за 9

месяцев 2024 года в России произошло более 96 тыс. ДТП, в которых погибли или получили травмы около 130 тыс. человек, представленных на рисунке 1.



Рисунок 1. Динамика показателей аварийности в России за девять месяцев в период с 2018 г. по 2024 г.

Источник: заимствовано из «Статистические данные о дорожно-транспортной аварийности в Российской Федерации за 9 месяцев 2024 г.» – URL: <https://media.mvd.ru/files/embed/8092246> (дата обращения: 18.09.2024)

Аварийность на дорогах по-прежнему является актуальной проблемой, а вероятность гибели пешеходов, велосипедистов и других уязвимых участников дорожного движения остается высокой. Для снижения аварийности и повышения безопасности дорожного движения необходимо проводить комплекс мероприятий. Важным направлением является использование современных технологий и интеллектуальных транспортных систем [2]. На сегодняшний день существуют мобильные приложения, которые информируют водителей о загруженности дорог, погодных условиях и опасных зонах на дорогах. Эти меры помогают уменьшить аварийность. Ключевым элементом остается анализ ДТП и мониторинг статистики в наиболее аварийных участках.

Анализ причин аварийности на дорогах осложняется множеством случайных факторов, влияющих на ДТП. Необходим комплексный подход, который включает улучшение инфраструктуры, повышение навыков водителей и изменение культуры вождения. Консолидация и анализ этих данных для получения выводов представляет собой сложную задачу.

Целью исследования является сравнение фактических показателей аварийности в Хабаровском крае за период с 01.01.2015 по 30.11.2023 гг. с результатами разработанной модели прогнозирования аварийности, путем ее обучения и апробации на основе рекуррентной нейронной сети с использованием методов машинного обучения.

Задачи настоящего исследования:

- проанализировать и подготовить фактические данные о количественных показателях ДТП в Хабаровском крае за период с 01.01.2015 по 30.11.2023 гг. для последующего использования при разработке модели прогнозирования;
- обучить и апробировать модель прогнозирования, которая будет воспроизводить общие тенденции и предсказывать пики количества ДТП, используя сравнительную характеристику;
- сравнить спрогнозированные значения разработанной модели с фактическими показателями ДТП в заданном периоде, определив степень точности и возможность ее дальнейшего применения и перспективы развития.

¹ Безопасность дорожного движения: Федеральный проект / утв. протоколом заседания проектного комитета по национальному проекту «Безопасные и качественные автомобильные дороги» от 20.12.2018. – URL: https://www.consultant.ru/document/cons_doc_LAW_319305/ (дата обращения: 12.02.2025).

² Статистические данные о дорожно-транспортной аварийности в Российской Федерации за 8 месяцев 2024 г. – URL: <https://media.mvd.ru/files/embed/8092246> (дата обращения: 18.09.2024).

Технологии искусственного интеллекта, такие как нейронные сети, генетические алгоритмы и алгоритмы кластеризации, в настоящее время активно применяются для выявления закономерностей в данных о дорожно-транспортных происшествиях и прогнозирования вероятности возникновения аварийных ситуаций. Использование машинного обучения позволяет адаптировать модели к конкретным условиям и повышает точность прогнозирования, что делает это направление исследований особенно перспективным для будущего развития. Этой теме посвящены многие публикации [1; 2; 4; 6; 11; 12; 13; 14; 15]. При работе над исследованием были рассмотрены различные методы и технологии, применяемые для анализа дорожно-транспортных происшествий, с использованием современных подходов на основе искусственного интеллекта.

Анализ признакового пространства

Аварийность во времени формируется случайным образом, но наблюдаются определенные статисти-

ческие закономерности. Во многом определяющим фактором формирования статистики ДТП является сезонность. Используя данные из открытых источников о ДТП, произошедших в Хабаровском крае за определенный временной диапазон (с 1 января 2015 г. по 30 ноября 2023 г.)³, можно проследить характерные особенности аварийности, выраженные в количественном распределении ДТП по месяцам года (рисунок 2) и с учетом сезонности (рисунок 3).

На основании гистограммы, представленной на рисунке 2, можно выделить некоторые тенденции и тренды:

- количество ДТП возрастает в тёплое время года, начиная с весны, достигая пика в летние и осенние месяцы, а затем снижается к зиме;
- наибольшее количество ДТП происходит в августе, сентябре и октябре более 1500 в месяц, в зимние время в январе, феврале – декабре – около 1000 случаев в месяц фиксируется минимальное количество ДТП.

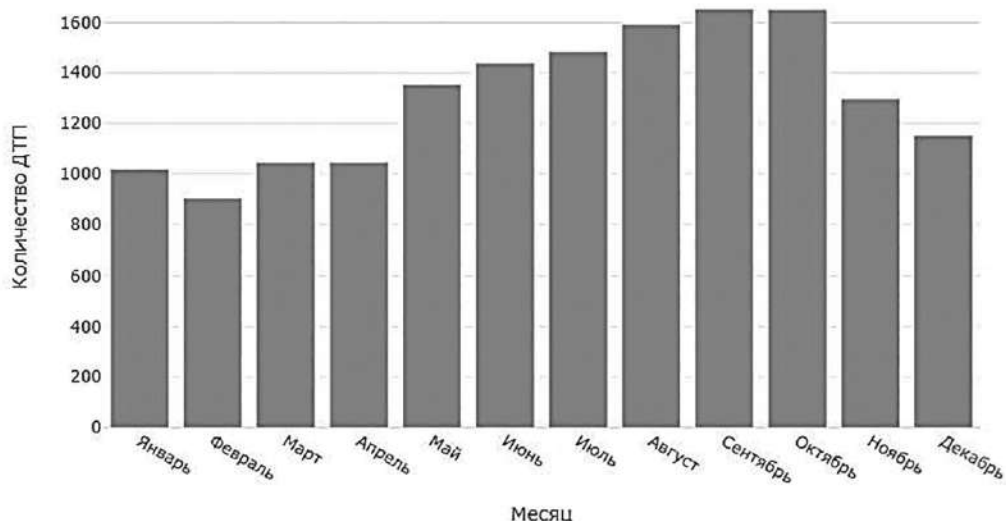


Рисунок 2. Распределение дорожно-транспортных происшествий по месяцам года в период с 01.01.2015 по 30.11.2023

Источник: разработано авторами на основе «Статистики ДТП по России» – URL: <https://rusdtp.ru/stat-dtp/> (дата обращения: 12.01.2025)

Распределение ДТП по месяцам имеет свои особенности:

- увеличение количества ДТП в марте и апреле по сравнению с зимним периодом, что может быть связано с улучшением погодных условий и увеличе-

нием активности на дорогах;

- июнь–октябрь: период максимальной активности ДТП, что может быть связано с увеличением числа путешествий, отпусков и активного автомобильного движения;

³ Дорожно-транспортные происшествия, произошедшие в Хабаровском крае в период с 1 января 2015г. по 30 ноября 2023г. – URL: <https://rusdtp.ru/stat-dtp/> (дата обращения: 12.01.2025).

– ноябрь и декабрь: снижение числа ДТП по сравнению с пиковыми месяцами, что, вероятно, связано со снижением активности водителей.

Гистограмма демонстрирует сезонные колебания числа ДТП. Летние и осенние месяцы являются наиболее опасными с точки зрения количества происшествий. Зимние месяцы относительно безопаснее, вероятно, из-за погодных условий и повышенной осторожности водителей.

На рисунке 3 представлена гистограмма, демонстрирующая распределение количества ДТП по вре-

менам года.

На гистограмме прослеживаются следующие тенденции:

- максимальное количество ДТП приходится на осенние месяцы (более 4000 случаев);
- лето также характеризуется высоким числом происшествий, сопоставимым с осенним периодом;
- весной зафиксировано умеренное количество ДТП – около 3500 случаев;
- минимальное число ДТП зафиксировано зимой (около 3200 случаев).

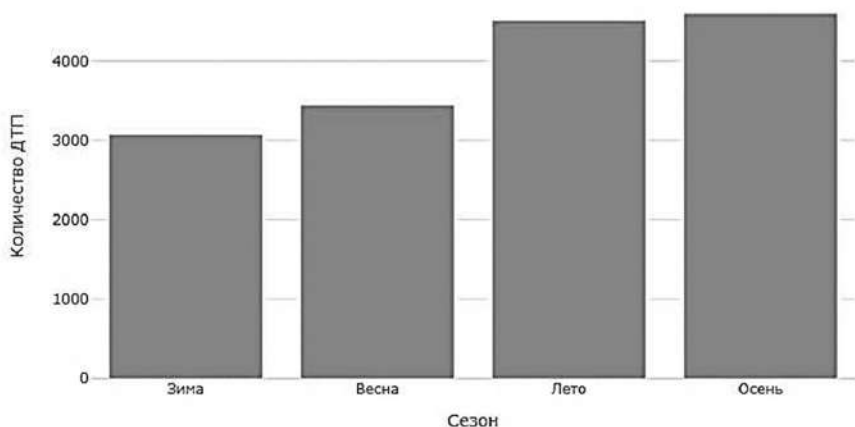


Рисунок 3. Распределение дорожно-транспортных происшествий по сезонам года в период с 01.01.2015 по 30.11.2023

Источник: разработано авторами на основе «Статистики ДТП по России» – URL: <https://rusdtp.ru/stat-dtp/> (дата обращения: 12.01.2025)

Гистограмма показывает ярко выраженную сезонную зависимость количества ДТП. Летний и осенний периоды являются наиболее опасными по аварийности, тогда как зимний период относительно безопаснее.

Анализ аварийности также включает изучение статистических данных о ДТП, что позволяет выявлять тенденции и отслеживать изменения в количественном показателе происшествий [6]. Также анализируются сезонные закономерности и влияние погоды. Исследование обстоятельств ДТП помогает оценить эффективность мер безопасности в дорожной инфраструктуре. Результаты анализа могут быть использованы для планирования и корректировки мероприятий по повышению безопасности на дорогах.

Принимая во внимание большой объем данных при проведении анализа аварийности с целью последующего прогнозирования, можно рассмотреть

внедрение интеллектуальных транспортных систем, основанных на применении таких методов машинного обучения, как нейронные сети⁴.

Современные технологии и методы машинного обучения обладают высокой эффективностью в прогнозировании благодаря способности обобщать сложные и скрытые зависимости в данных [8]. Они позволяют обрабатывать большие объемы информации, выявлять неочевидные связи между переменными, что является преимуществом по сравнению с рутинными методами статистического анализа. Адаптивность алгоритмов к новым данным способствует улучшению точности последующих прогнозов. Методы машинного обучения выявляют взаимодействия между факторами, влияющими на результаты, и позволяют моделировать сложные системы с помощью различных моделей, таких как нейронные сети [4]. Применение

⁴ Жигалов К. Ю. Использование современных подходов в области машинного обучения для предиктивного анализа потенциальных отказов на основе исторических данных и показателей состояния оборудования // Актуальные вопросы науки и образования: Сборник материалов X Международной научно-практической конференции, Москва, 08 июля 2024 года. – М., 2024. – С. 195–202. – EDN: JUQKQW.

машинного обучения в анализе дорожной аварийности способствует эффективной обработке данных и принятию обоснованных решений для повышения безопасности на дорогах. Это значительно улучшает точность прогнозов и позволяет принимать более обоснованные решения при планировании мероприятий, направленных на повышение уровня безопасности.

Модель прогнозирования на основе статистических данных

В основе расчета были использованы данные из открытых источников об аварийности в Хабаровском крае и общие характеристики за определенный временной диапазон с 1 января 2015 г. по 30 ноября 2023 г.

Для обучения прогностической нейросетевой модели использовались регулярные временные ряды с параметрами ДТП. В качестве признакового пространства модели применялись следующие численные характеристики ДТП, рассчитанные с интервалом в 1 день:

- количество ДТП;
- участники (общее количество участников ДТП);
- пострадавшие (число пострадавших);
- смертность (число смертельных случаев);
- серьезность ДТП;
- параметр, характеризующий недельную сезонность;
- параметр, характеризующий годовую сезонность;
- состояние дорожного покрытия, вызванное погодными условиями;
- дневной или ночной режим.

Итоговые визуализированные данные, полученные с помощью нейронной сети, могут служить основой для анализа произошедших ДТП в заданный период. Для анализа признакового пространства разрабатываемых моделей использовались диаграммы виолончельного типа (violinplots), представленные на рисунке 4.

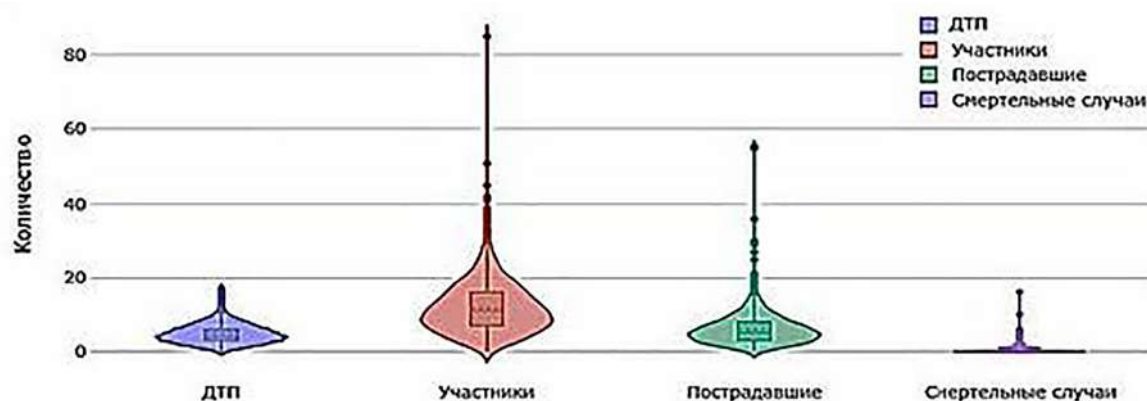


Рисунок 4. Виолончельный график взаимного распределения параметров дорожно-транспортных происшествий, рассчитанных как сумма за 1 день

Источник: разработано авторами

На представленном графике изображены распределения различных параметров, связанных с ДТП за один день. Данные графика:

- количество ДТП: диаграмма узкая, что указывает на малую вариативность в данных; среднее значение находится в районе 2–5;
- количество участников: распределение более широкое, с длинными хвостами, что указывает на значительные отклонения; большая часть значений лежит в пределах до 20;
- количество пострадавших: распределение также довольно широкое, с длинным хвостом, направленным вверх; большинство значений сосредоточено в области до 10 пострадавших, но есть редкие случаи с большим количеством;

- количество смертельных случаев: диаграмма очень узкая, что указывает на низкую вариативность; число смертельных случаев в основном близко к 0, что согласуется с редкостью таких событий.

На основании анализа полученного графика можно сделать следующие выводы:

- большинство ДТП за один день связано с малым количеством пострадавших участников и смертельных случаев;
- категории «Участники» и «Пострадавшие» демонстрируют наибольшую вариативность и плотность данных;
- смертельные случаи встречаются редко, что характеризуется узкой и низкой формой диаграммы.

На рисунке 5 представлено попарное распределение параметров ДТП.

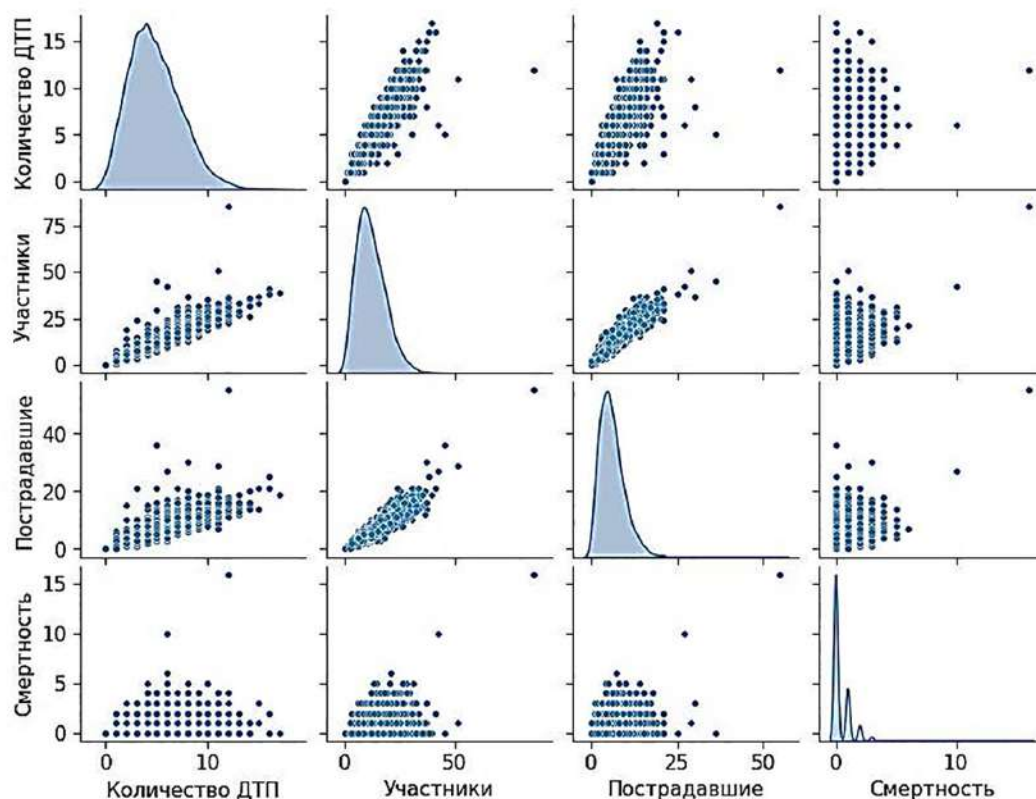


Рисунок 5. Графики взаимного распределения параметров ДТП, рассчитанные с периодичностью в 1 день
Источники: разработано авторами

Такие графики позволяют выполнять совместный анализ рассматриваемых признаков, с целью определения корреляционных зависимостей и исследовать плотность распределения отдельных параметров.

График построен для следующих параметров:

- количество ДТП (за день),
- участники (общее количество участников ДТП за день),
- пострадавшие (число пострадавших за день),
- смертность (число смертельных случаев за день).

Каждый элемент матрицы представляет собой зависимость двух переменных. Например, график в первой строке и втором столбце показывает взаимосвязь между количеством ДТП и участниками. Диагональные элементы содержат графики плотности для распределения каждого признака. Элементы ниже диагонали повторяют информацию из верхней части матрицы, но отображаются для других осей.

На графиках видна чёткая линейная зависимость – с увеличением числа ДТП увеличивается и количест-

во участников. Большинство же значений сосредоточено в диапазоне до 10 ДТП и до 50 участников.

Прослеживается прямая зависимость – чем больше ДТП, тем больше пострадавших. Смертельные случаи остаются редким событием, независимо от числа ДТП. Зависимость значения «участники» от значения «пострадавшие» ближе к линейной – при увеличении числа участников растёт число пострадавших. Зависимость «смертности» от «пострадавших» является закономерной. С увеличением числа пострадавших увеличивается и число смертельных случаев.

На основе анализа попарного распределения можно сделать следующие выводы:

- увеличение количества ДТП ведёт к увеличению числа участников и пострадавших;
- число смертельных случаев значительно реже, но оно также коррелирует с другими признаками, особенно с числом пострадавших;
- распределение параметров неравномерно, основная масса данных сосредоточена около низких значений.

Подготовка признакового пространства с учетом кодирования исходных данных

Для подготовки исходных данных, их необходимо классифицировать и преобразовать, учитывая степень серьезности ДТП. Используя кодирование, заменяем категории числами: легкая – 1, тяжелая – 2, смертельная – 3.

$$\begin{aligned} \text{day_sin} &= \sin\left(\frac{2\pi \cdot \text{day_of_week}}{7}\right) & \text{month_sin} &= \sin\left(\frac{2\pi \cdot \text{day_of_year}}{365}\right) \\ \text{day_cos} &= \cos\left(\frac{2\pi \cdot \text{day_of_week}}{7}\right) & \text{month_cos} &= \cos\left(\frac{2\pi \cdot \text{day_of_year}}{365}\right). \end{aligned} \quad (1)$$

Учёт состояния дороги выполнялся с применением «dummy-кодирования» для преобразования категориальных данных в числовые. В частности, к таким параметрам относилось сухое или мокрое состояние дорожного покрытия, светлое или темное время суток. Таким образом, по каждой категории создавался отдельный признак и при наличии соответствующего условия в ДТП, величина такого признака принимала значение 1, а при его отсутствии – 0.

Разработка и обучение авторегрессионной модели

Для моделирования и последующего прогнозирования использовалась LSTM (LongShort-TermMemory) сеть – это тип рекуррентной нейронной сети (RNN), специально разработанной для работы с последовательными данными и временными рядами [14]. Данный тип сетей при помощи специальных ячеек памяти сохраняет информацию, предотвращая проблему исчезающего градиента.

Данные, учитывающие особенности недельной и годовой сезонности, можно преобразовать путём использования тригонометрических формул синуса и косинуса, позволяющих создать плавные переходы между днями недели, что позволит разрабатываемой модели лучше уловить структуру временного ряда.

Из преимуществ использования сетей типа LSTM при прогнозировании временных рядов можно выделить следующие:

- LSTM хорошо справляется с задачами, где необходимо учитывать информацию в любые промежутки времени и наблюдаются долгосрочные зависимости;
- LSTM сети могут самостоятельно решать, какую информацию сохранять исходя из заданных параметров при вводе и дальнейшем использовании входных данных, (входной гейт) и определять, какую информацию использовать для предсказаний (выходной гейт), определять лишнюю информацию, не используемую при последующих действиях [11].

Моделирование производилось с применением авторегрессионного подхода [12]. Данный метод основан на предположении, что текущее значение временного ряда можно выразить через его предыдущие значения.

На рисунке 6 представлена визуализация работы авторегрессионной модели.

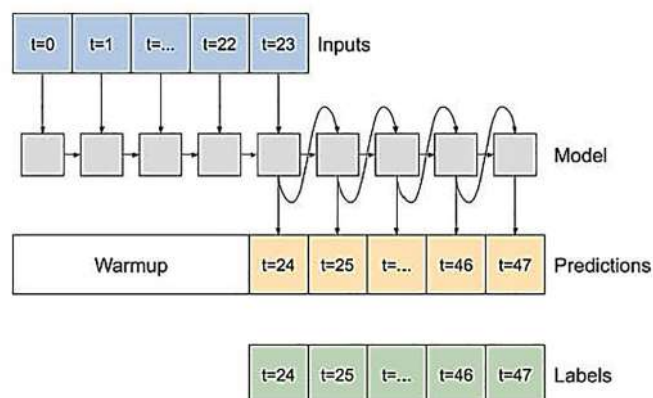


Рисунок 6. Визуализация работы авторегрессионной модели

Источник: заимствовано из Lastrico, R., Macciò, S., Carfi, A., Traverso, P., Mastrogiovanni, F. (2024). Estimation of Kidney's Blood Vessels Deformations for Robot-Assisted Surgery. In: Lee, SG., An, J., Chong, N.Y., Strand, M., Kim, J.H. (eds) Intelligent Autonomous Systems 18. IAS 2023. Lecture Notes in Networks and Systems, Vol. 794. Springer, Cham. – https://doi.org/10.1007/978-3-031-44981-9_35

Выходные данные каждой модели могут быть возвращены в себя же на каждом шаге с целью их использования для последующих прогнозов.

Согласно принципам работы авторегрессионной модели входные данные выборочной совокупности необходимо разделить на обучающую часть – (70 %), валидационную (для промежуточного контроля полученных данных) – (20 %) и тестовую (для итоговой оценки модели) – (10 %) [13].

Правильное разделение данных является критиче-

ски важным шагом в процессе прогнозирования временных рядов, так как оно обеспечивает надежность и достоверность результатов. Временные ряды имеют естественный порядок, и случайное разделение может привести к утрате важной информации о временной зависимости [5]. Последовательное разделение позволяет правильно оценить производительность модели на данных, которые она не «видела» во время обучения.

Важным элементом разработанной модели являлось использование функции потерь HuberLoss [15]

$$L_{\delta}(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{если } |y - \hat{y}| \leq \delta \\ \delta \cdot |y - \hat{y}| - \frac{1}{2}\delta^2 & \text{если } |y - \hat{y}| > \delta, \end{cases} \quad (2)$$

где

y – истинное значение;

$\hat{y}$ – предсказанное значение;

δ – порог, который определяет, когда использовать квадратичную или линейную ошибку.

Данная функция используется для обучения нейросетей в задачах регрессии, когда важно учитывать как точность, так и устойчивость к выбросам. Она объединяет в себе преимущества двух популярных функций: среднеквадратичной ошибки (MSE) и абсолютной ошибки (MAE).

Обучение и апробация модели прогнозирования

Рассмотрим примеры прогнозных значений по количеству ДТП на обучающем наборе данных. Производилось обучение модели в течении 100 эпох. Эпоха в машинном обучении – это один полный проход по

всему обучающему набору данных во время тренировки модели [1]. В процессе обучения модель обновляет свои параметры методом оптимизации полученных результатов.

Пример 1. Исходные данные, принимаемые обученной моделью на входе, являлись данные по ДТП из обучающего набора с 1 января по 31 декабря 2018 г. Целью являлся прогноз аварийности на январь 2019 г. на 30 дней.

График, представленный на рисунке 7, иллюстрирует процесс сравнения фактических данных со значениями, являющимися результатом прогнозирования, полученными на выходе модели.

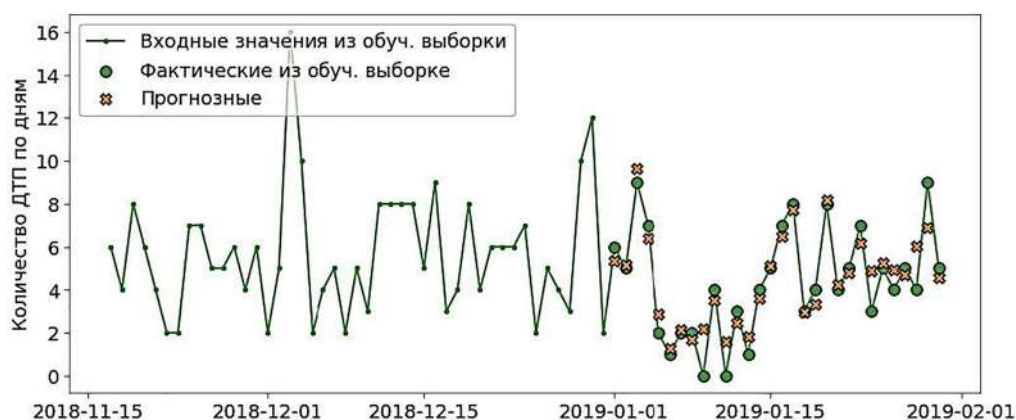


Рисунок 7. Временной ряд количества ДТП по дням заданного периода обучения модели и последующего прогнозирования с визуальным сравнением фактических и спрогнозированных данных

Источник: разработано авторами

По оси Х отложена временная шкала, охватывающая период с ноября 2018 года по февраль 2019 года. По оси Y указано количество ДТП за каждый день.

Зелёная линия с маркерами (точки): отображает входные значения, взятые из обучающей выборки. Эта линия соединяет фактические данные, используемые для построения модели. Зеленые круги (маркеры): показывают фактические данные из обучающей выборки, отображающие истинное количество ДТП. Оранжевые кресты (маркеры): представляют прогнозные значения модели для соответствующих дней.

На графике наблюдается согласованность между фактическими (зеленые круги) и прогнозными значениями (оранжевые кресты). Однако есть расхождения, особенно в периоде резких пиков, например, в начале декабря 2018 года со всплеском ДТП. Прогнозы в целом следуют за тенденцией временного ряда. График показывает, как модель прогнозирует

количество ДТП по временным данным. Прогнозы, в основном, близки к фактическим, что свидетельствует об удовлетворительном обучении модели, но существуют перспективы улучшения точности при экстремальных значениях.

Пример 2. На вход обученная модель принимает данные по ДТП из обучающего набора с 1 октября 2022 г. по 30 сентября 2023 г. Ожидаемый результат – прогноз аварийности на октябрь 2023 г. сразу на 30 дней. Прогнозы делаются для данных, которые не были задействованы в обучении. Границей данных обучающей и тестовой выборками на графике является стык зеленой и синей линий. График, представленный на рисунке 8, иллюстрирует процесс проверки возможностей прогнозирования модели на новых данных.

По оси Х отображена временная шкала, охватывающая период с декабря 2022 года по февраль 2023 года. По оси Y указано количество ДТП за день.

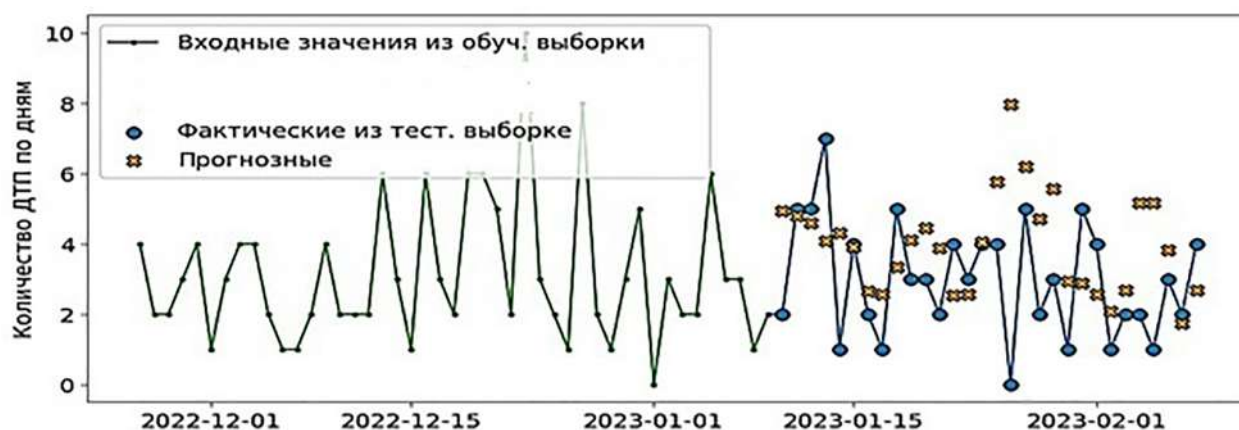


Рисунок 8. Временной ряд количества ДТП по дням, разделённый на обучающую и тестовую выборки
Источник: разработано авторами

Зелёная линия с маркерами (точки): обозначает входные значения из обучающей выборки. Это фактические данные, на которых обучалась модель. Синие маркеры (круги) характеризуют фактические значения тестовой выборки. Это данные, которые не участвовали в обучении модели и используются для проверки её качества. Оранжевые кресты представляют прогнозные значения модели для тестовой выборки. Синяя линия отображает динамику фактических данных тестовой выборки (соседние точки соединены для наглядности).

Данная модель воспроизводит общие тенденции и предсказывает пики количества ДТП. Прогнозные

значения имеют меньший разброс, чем фактические данные, что может свидетельствовать о сглаживании экстремальных значений моделью.

Можно сделать вывод, что модель, обученная на данных из обучающей выборки (зелёная линия), способна предсказывать количество ДТП на данных тестовой выборки (синие маркеры). Значения, полученные в результате прогнозирования, в целом согласуются с фактическими, но точность модели может снижаться в периодах с резкими изменениями или пиками. Для повышения точности можно применить дополнительные признаки или адаптацию модели к тестовой выборке.

Выводы и направления дальнейшего совершенствования

Полученные результаты прогнозирования ДТП в Хабаровском крае на основе фактических данных с использованием авторегрессионной нейросетевой модели можно признать удовлетворительными. Это направление имеет перспективы для улучшения качества и точности прогнозирования, что может помочь в профилактике и снижении аварийности на автомобильных дорогах с использованием доступных онлайн и офлайн ресурсов [7]. Последующее совершенствование модели прогнозирования может заключаться в применении большего объема информации в исходных данных, с более широким спектром анализируемых признаков (координаты места происшествий, гендерная принадлежность водителя, марка транспортного средства и др.).

Представленная в исследовании модель прогнози-

рования может быть адаптирована для предсказания других параметров ДТП, например таких, как:

- прогнозирование количества ДТП, участников и пострадавших;
- предсказание видов ДТП (легкие, тяжелые, со смертельным исходом).

Дальнейшая работа потребует увеличения объема фактических данных с учетом дополнительных обстоятельств происшествий и более подробного рассмотрения характеристик их участников.

Результаты исследования в виде полученной модели прогнозирования могут быть адаптированы для прогнозирования количества и видов ДТП, участников и пострадавших. Применение результатов исследования может быть направлено на улучшение качества профилактики дорожной аварийности и предотвращения аварийных ситуаций, используя данные модели прогнозирования.

Литература

1. Велихов П. Машинное обучение для понимания естественного языка // Открытые системы СУБД. – 2016. – № 1. – С. 18–21. – EDN: VQCODX.
2. Гущина П. В., Шакирзянова И. Ф., Зюрина О. А. Интеллектуальные транспортные системы: элементы и преимущества // Логистика – Евразийский мост. Материалы XIX Международной научно-практической конференции. Красноярск, 24–28 апреля 2024. – С. 70–75. – EDN: SNSONE.
3. Дорожно-транспортный травматизм. Вред, причиняемый во время ДТП / А. И. Марковнина [и др.] // Будущее технической науки. Сборник материалов XIX Всероссийской молодежной научно-технической конференции. Нижний Новгород, 08 октября 2020. – С. 190–191. – EDN: QFYGBA.
4. Звягин Л. С. Особенности моделирования в различных типах нейросетей // Мягкие измерения и вычисления. – 2022. – Т. 52, № 3. – С. 43–54. – <https://doi.org/10.36871/2618-9976.2022.03.005>. – EDN: TMFOXC.
5. Канторович Г. Г. Анализ временных рядов // Экономический журнал Высшей школы экономики. – 2002. – Т. 6, № 2. – С. 251–273. – EDN: ZEKQFV.
6. Кущенко Л. Е. Прогнозирование ДТП как один из способов снижения смертности // Мир транспорта и технологических машин. – 2023. – № 2(81). – С. 67–73. – [https://doi.org/10.33979/2073-7432-2023-2\(81\)-67-73](https://doi.org/10.33979/2073-7432-2023-2(81)-67-73). – EDN: MNZHMS.
7. Обобщённая схема анализа данных и прогнозирования аварийных ситуаций на автомобильных дорогах / Д. О. Бобынцев [и др.] // Естественные и технические науки. – 2021. – № 12 (163). – С. 285–288. – EDN: NTGEKA.
8. Особенности методов машинного обучения / И. Я. Львович [и др.] // Информационные технологии. – 2020. – Т. 26, № 9. – С. 499–506. – <https://doi.org/10.17587/it.26.499-506>. – EDN: EIZONB.
9. Рыбин А. Л. Парадигмы дорожного движения // Автомобильные дороги. – 2020. – № 12 (1069). – С. 133–137. – EDN: ZDUVPM.
10. Шпалтаков В. П. Теоретические основы экономики транспортной системы // Инновационная экономика и общество. – 2020. – № 1 (27). – С. 2–14. – EDN: KPFPFW.
11. Alvarez-Rodriguez U., et al. (2021). Evolutionary dynamics of higher-order interactions in social networks. *Nature Human Behaviour*, Vol. 5, No. 5, pp. 586–595. – <https://doi.org/10.1038/s41562-020-01024-1>. (In Eng.).
12. Franco Basso, et al. (2021) A deep learning approach for real-time crash prediction using vehicle-by-vehicle data. *Accident Analysis & Prevention*, Vol. 162. – <https://doi.org/10.1016/j.aap.2021.106409>. (In Eng.).
13. Lastrico R., et al. (2024). Estimation of Kidney's Blood Vessels Deformations for Robot-Assisted Surgery. In: Lee SG., An J., Chong N. Y., Strand M., Kim J. H. (eds) *Intelligent Autonomous Systems 18. IAS 2023. Lecture Notes in Networks and Systems*, Vol. 794. – https://doi.org/10.1007/978-3-031-44981-9_35. (In Eng.).
14. Li P., Abdel-Aty M., Yuan, J. (2020) Real-Time Crash Risk Prediction on Arterials Based on LSTM-CNN. *Accident Analysis & Prevention*, Vol. 135. – <https://doi.org/10.1016/j.aap.2019.105371>. (In Eng.).

15. Liu Q., et al. (2024) Multi-task Support Vector Machine Classifier with Generalized Huber Loss. *Journal of Classification*. – <https://doi.org/10.1007/s00357-024-09488-w>. (In Eng.).

References

1. Velikhov, P. (2016) [Machine Learning for Natural Language Understanding]. *Otkrytyye sistemy SUBD* [Open DBMS Systems]. Vol. 1, pp. 18–21. (In Russ.).
2. Gushchina, P. V., Shakirzyanova, I. F., Zyurina, O. A. (2024) [Intellectual transportation systems: elements and advantages]. *Logistika – Yevraziyskiy most. Materialy XIX Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Logistics – Eurasian Bridge. Materials of XIX International Scientific and Practical Conference]. Krasnoyarsk, pp. 70–75. (In Russ.).
3. Markovnina, A. I., et al. (2020) [Road traffic injuries. Harm caused during road traffic accidents]. *Budushcheye tekhnicheskoy nauki. Sbornik materialov XIX Vserossiyskoy molodezhnoy nauchno-tekhnicheskoy konferentsii* [Future of technical science. Proceedings of the XIX All-Russian Youth Scientific and Technical Conference]. Nizhny Novgorod, October 08, pp. 190–191. (In Russ.).
4. Zvyagin, L. S. (2022) [Peculiarities of modeling in different types of neural networks]. *Myagkiye izmereniya i vychisleniya* [Soft Measurements and Computations]. Vol. 52. No. 3, pp. 43–54. – <https://doi.org/10.36871/2618-9976.2022.03.005>. (In Russ.).
5. Kantorovich, G. G. (2002) [Analysis of time series] *Ekonomicheskiy zhurnal Vysshey shkoly ekonomiki* [Economic Journal of the Higher School of Economics]. Vol. 6, No. 2, pp. 251–273. (In Russ.).
6. Kushchenko, L. E. (2023) [Prediction of traffic accidents as one of the ways to reduce mortality]. *Mir transporta i tekhnologicheskikh mashin* [World of transport and technological machines]. Vol. 2(81), pp. 67–73. – [https://doi.org/10.33979/2073-7432-2023-2\(81\)-67-73](https://doi.org/10.33979/2073-7432-2023-2(81)-67-73). (In Russ.).
7. Bobintsev, D. O. (2021) [Generalized scheme of data analysis and forecasting of emergency situations on highways]. *Yestestvennyye i tekhnicheskiye nauki* [Natural and Technical Sciences]. Vol. 12, pp. 285–288. (In Russ.).
8. Lvovich, I. Y. (2020) [Features of machine learning methods]. *Informatsionnyye tekhnologii* [Information technologies]. Vol. 26. No. 9, pp. 499–506. – <https://doi.org/10.17587/it.26.499-506>. (In Russ.).
9. Rybin, A. L. (2020) [Paradigms of road traffic]. *Avtomobil'nyye dorogi* [Automobile roads]. Vol. 1069, pp. 133–137. (In Russ.).
10. Shpaltakov, V. P. (2021) [Theoretical foundations of the economy of the transport system]. *Innovatsionnaya ekonomika i obshchestvo* [Innovative economy and society]. Vol. 1 (27), pp. 2–14. (In Russ.).
11. Alvarez-Rodriguez, U., et al. (2021) Evolutionary dynamics of higher-order interactions in social networks. *Nature Human Behaviour*, Vol. 5, No. 5, pp. 586–595. – <https://doi.org/10.1038/s41562-020-01024-1>. (In Eng.).
12. Franco Basso, et al. (2021) A deep learning approach for real-time crash prediction using vehicle-by-vehicle data. *Accident Analysis & Prevention*, Vol. 162. – <https://doi.org/10.1016/j.aap.2021.106409>. (In Eng.).
13. Lastrico, R., et al. (2024) Estimation of Kidney's Blood Vessels Deformations for Robot-Assisted Surgery. In: Lee, S.G., An, J., Chong, N.Y., Strand, M., Kim, J.H. (eds) *Intelligent Autonomous Systems 18. IAS 2023. Lecture Notes in Networks and Systems*, Vol. 794. Springer, Cham. – https://doi.org/10.1007/978-3-031-44981-9_35. (In Eng.).
14. Li, P., Abdel-Aty, M., Yuan, J. (2020) Real-Time Crash Risk Prediction on Arterials Based on LSTM-CNN. *Accident Analysis & Prevention*, Vol. 135. – <https://doi.org/10.1016/j.aap.2019.105371>. (In Eng.).
15. Liu, Q., et al. (2024) Multi-task Support Vector Machine Classifier with Generalized Huber Loss. *Journal of Classification* – <https://doi.org/10.1007/s00357-024-09488-w>. (In Eng.).

Информация об авторах:

Егор Викторович Шохирев, аспирант, научная специальность 2.9.5 Эксплуатация автомобильного транспорта, Тихоокеанский государственный университет, Хабаровск, Россия

ORCID iD: 0009-0002-5290-8064

e-mail: shohirev_kna@mail.ru

Павел Павлович Володькин, доктор технических наук, профессор, руководитель высшей школы «Транспортные системы и технологии», Тихоокеанский государственный университет, Хабаровск, Россия

ORCID iD: 0000-0002-2731-2069

e-mail: 004167@togudv.ru

Владимир Александрович Лазарев, кандидат технических наук, доцент высшей школы «Транспортные системы и технологии», Тихоокеанский государственный университет, Хабаровск, Россия

ORCID iD: 0000-0003-3768-7793

e-mail: 000136@togudv.ru

Александр Викторович Константинов, научный сотрудник, Институт горного дела Дальневосточного отделения Российской академии наук, Хабаровск, Россия

ORCID iD: 0000-0001-6481-292X, **ResearcherID:** AAL-6888-2021, **ScopusID:** 57204216063

e-mail: alex-sdt@yandex.ru

Вклад соавторов:

Шохирев Е. В. – реализация задач исследования, обзор литературных источников, разработка проведения исследования, формулировка основных принципов разработки и обучения авторегрессионной модели прогнозирования, анализ и обсуждение результатов, формулировка выводов.

Володькин П. П. – постановка цели и задачи исследования, анализ и обсуждение результатов, формулировка выводов, научная редакция статьи.

Лазарев В. А. – обзор литературных источников, анализ и обсуждение результатов, формулировка выводов, подготовка и научная редакция статьи.

Константинов А. В. – формулировка основных принципов разработки и обучения авторегрессионной модели прогнозирования, подготовка исходных данных с учетом кодирования.

Конфликт интересов отсутствует.

Статья поступила в редакцию: 03.03.2025; принята в печать: 22.05.2025.

Авторы прочитали и одобрили окончательный вариант рукописи.

Information about the authors:

Egor Viktorovich Shokhirev, postgraduate student, scientific specialty 2.9.5 Operation of Road Transport, Pacific National University, Khabarovsk, Russia

ORCID iD: 0009-0002-5290-8064

e-mail: shohirev_kna@mail.ru

Pavel Pavlovich Volodkin, Doctor of Technical Sciences, Professor, Head of the higher school «Transport Systems and Technologies», Pacific National University, Khabarovsk, Russia

ORCID iD: 0000-0002-2731-2069

e-mail: 004167@togudv.ru

Vladimir Aleksandrovich Lazarev, Candidate of Technical Sciences, Associate Professor of the higher school «Transport Systems and Technologies», Pacific National University, Khabarovsk, Russia

ORCID iD: 0000-0003-3768-7793

e-mail: 000136@togudv.ru

Aleksander Viktorovich Konstantinov, researcher, Mining Institute of the Far Eastern Branch of the Russian Academy of Sciences, Khabarovsk, Russia

ORCID iD: 0000-0001-6481-292X, **ResearcherID:** AAL-6888-2021, **ScopusID:** 57204216063

e-mail: alex-sdt@yandex.ru

Contribution of the authors:

Shokhirev E. V. – implementation of study objectives, review of literature sources, design of study conduct, formulation of basic principles for the development and training of an autoregressive prediction model, analysis and

discussion of results, formulation of conclusions.

Volodkin P. P. – setting the goal and objectives of the study, analysis and discussion of the results, formulation of conclusions, scientific version of the article.

Lazarev V. A. – review of literature, analysis and discussion of results, formulation of conclusions, preparation and scientific revision of the article.

Konstantinov A. V. – formulation of basic principles of development of autoregressive forecasting model training, preparation of initial data taking into account coding.

There is no conflict of interest.

The paper was submitted: 03.03.2025.

Accepted for publication: 22.05.2025.

The authors have read and approved the final manuscript.