

ТРАНСПОРТ

УДК 004.89

<https://doi.org/10.25198/2077-7175-2022-6-69>

ИНТЕЛЛЕКТУАЛЬНАЯ МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ИНТЕНСИВНОСТИ ДВИЖЕНИЯ ТРАНСПОРТНЫХ СРЕДСТВ НА ПЕРЕКРЕСТКЕ

И. П. Болодурина¹, Л. М. Анциферова², Л. С. Гришина³

Оренбургский государственный университет, Оренбург, Россия

¹ e-mail: ipbolodurina@yandex.ru² e-mail: antsiferova_68@mail.ru³ e-mail: grishina_ls@inbox.ru

Аннотация. Активное развитие парка транспортных средств способствует увеличению загруженности транспортных сетей и требует построения новых систем определения мест перегрузки, а также разработки способов разрешения сложившихся проблем. Особое место в данной области занимает ранняя оценка состояния транспортных сетей, которая позволяет своевременно принять меры по их модификации. В связи с этим, исследование в области прогнозирования основных показателей транспортного потока с целью выявления перегрузок транспортных сетей приобрели высокую актуальность. Использование технологий интеллектуального анализа данных, включая методы многомерной линейной регрессии, позволяет построить прогнозные модели для исследования характеристик транспортного потока. **Цель исследования.** Построить модель прогнозирования интенсивности движения на перекрестке методами машинного обучения для эффективного принятия решений при управлении транспортным потоком. Научная новизна данного исследования заключается в разработке модели многомерной регрессии для прогнозирования интенсивности движения с учетом контроля переобучения на основе открытых данных о наблюдениях за количеством транспортных средств на четырех различных участках перекрестка. **Методы.** Для прогнозирования интенсивности движения транспортного потока использована модель многомерной линейной регрессии. В качестве методов повышения эффективности прогнозирования использованы подходы к регуляризации Ridge, Lasso и ElasticNet. **Результаты.** Построена модель прогнозирования интенсивности движения транспортного потока. В рамках экспериментального исследования проведена оценка эффективности применения методов регуляризации, а также сравнительный анализ точности прогнозирования модели на основе исходных и нормализованных данных. **Заключение.** Наименьшую среднеквадратичную ошибку на тестовых данных показала модель многомерной линейной регрессии с регуляризацией Ridge. Разработанная модель позволяет спрогнозировать количество транспортных средств, проходящих в единицу времени. Полученные результаты позволяют заранее выявлять места транспортной сети с повышенной нагрузкой для оптимизации движения транспортных средств на различных перекрестках УДС. Результаты проведенного исследования имеют высокую практическую значимость, так как могут быть внедрены в существующие системы управления дорожным движением для эффективного принятия решений при управлении транспортным потоком. Направление будущих исследований включает в себя практическую апробацию прогнозной модели в реальных условиях, а также рассмотрение ансамблевых методов машинного обучения для повышения точности при прогнозировании основных показателей транспортного потока.

Ключевые слова: транспортный поток, интенсивность движения, системы управления дорожным движением, методы машинного обучения, многомерная линейная регрессия, регуляризация.

Благодарности: работа выполнена при поддержке гранта Российского фонда фундаментальных исследований (№ 20-07-01065 «А»), а также стипендии Президента Российской Федерации для молодых ученых и аспирантов (№ СП-3652.2021.5).

Для цитирования: Болодурина И. П., Анциферова Л. М., Гришина Л. С. Интеллектуальная модель прогнозирования интенсивности движения транспортных средств на перекрестке // Интеллект. Инновации. Инвестиции. – 2022. – № 6. – С. 69–78, <https://doi.org/10.25198/2077-7175-2022-6-69>.

INTELLIGENT MODEL FOR PREDICTING THE TRAFFIC INTENSITY OF VEHICLES AT THE INTERSECTION

I. P. Bolodurina¹, L. M. Antsiferova², L. S. Grishina³

Orenburg State University, Orenburg, Russia

¹e-mail: ipbolodurina@yandex.ru;

²e-mail: antsiferova_68@mail.ru;

³e-mail: grishina_ls@inbox.ru

Abstract. The active development of the fleet of vehicles contributes to an increase in the congestion of transport networks and requires the construction of new systems for determining the places of congestion, as well as the development of ways to solve the existing problems. A special place in this area is occupied by an early assessment of the condition of transport networks, which allows timely measures to be taken to modify them. In this regard, research in the field of forecasting the main indicators of traffic flow in order to identify congestion of transport networks has become highly relevant. The use of data mining technologies, including methods of multidimensional linear regression, will make it possible to build predictive models for studying the characteristics of traffic flow. **Aim.** To build a model for predicting traffic intensity at an intersection using machine learning methods for effective decision-making when managing traffic flow. The scientific novelty of this study lies in the development of a multidimensional regression model for predicting traffic intensity, considering the control of retraining based on open data on observations of the number of vehicles at four different sections of the intersection. **Methods.** A multidimensional linear regression model was used to predict the intensity of traffic flow. Ridge, Lasso and ElasticNet regularization approaches were used as methods to improve the forecasting efficiency. **Results.** A model for predicting the intensity of traffic flow is constructed. As part of the experimental study, an assessment of the effectiveness of the use of regularization methods was carried out, as well as a comparative analysis of the accuracy of forecasting the model based on the initial and normalized data. **Conclusion.** The smallest root-mean-square error on the test data was shown by the multidimensional linear regression model with Ridge regularization. The developed model makes it possible to predict the number of vehicles passing per unit of time with an average quadratic error equal to 0.638. The obtained results of the study will allow early diagnostics of the occurrence of congestion of transport networks to optimize the movement of vehicles at various intersections of the UDS. The results of the study are of high practical importance, as they can be implemented into existing traffic management systems for effective decision-making in traffic flow management. The direction of future research includes the practical testing of the predictive model in real conditions, as well as the consideration of ensemble methods of machine learning to improve the accuracy in predicting the main indicators of the traffic flow.

Keywords: traffic flow, traffic intensity, traffic management systems, machine learning methods, multidimensional linear regression, regularization.

Acknowledgements: the work was supported by a grant from the Russian Foundation for Basic Research (No. 20-07-01065 “A”), as well as a scholarship from the President of the Russian Federation for young scientists and postgraduates (No. SP-3652.2021.5).

Cite as: Bolodurina, I. P., Antsiferova, L. M., Grishina, L. S. (2022) [Intelligent model for predicting the traffic intensity of vehicles at the intersection]. *Intellekt. Innovacii. Investicii* [Intellect. Innovations. Investments]. Vol. 6, pp. 69–78, <https://doi.org/10.25198/2077-7175-2022-6-69>.

Введение

Активное развитие парка транспортных средств способствует увеличению загруженности транспортных сетей и требует построения новых систем выявления мест перегрузки, а также разработки способов разрешения сложившихся проблем. Особое место в данной области занимает ранняя оценка состояния транспортных сетей, которая позволяет своевременно принять меры по их модификации. В связи с этим, исследования в области прогнозирования основных показателей транспортного потока с целью раннего выявления перегрузок транспортных сетей приобрели высо-

кую актуальность [13]. Использование технологий интеллектуального анализа данных, включая методы многомерной линейной регрессии, позволят построить прогнозные модели для исследования характеристик транспортного потока. При этом в настоящий момент разработаны лишь алгоритмы решения частных задач оценки и прогнозирования состояния транспортной сети, которые сложно адаптировать под решение новых или даже схожих задач.

С развитием новых технологий сбора данных с датчиков транспортных средств открылись возможности анализа больших объемов накопленных

данных о транспортном потоке [1, 9]. Прогнозирование транспортных заторов во все большей степени зависит от сбора данных, таких как информация о движении транспорта, скорости транспортных средств, общего времени в пути, плотности движения, погоде, авариях и т. д. [8]. Однако традиционные модели, такие как интегрированные модели авторегрессии скользящего среднего для временных рядов (ARIMA) [14] или методы опорных векторов (SVM) [7], не всегда могут хорошо адаптироваться к новым условиям.

Целью данной работы является построение модели прогнозирования интенсивности движения на перекрестке методами машинного обучения для эффективного принятия решений при управлении транспортным потоком. Научная новизна данного исследования заключается в построении модели многомерной регрессии для прогнозирования интенсивности движения с учетом контроля переобучения на основе открытых данных о наблюдениях за количеством транспортных средств на четырех различных участках перекрестка.

Обзор исследований

Исследованиями в области прогнозирования основных показателей транспортного потока занимаются ученые по всему миру.

Так, автор статьи [5] описывает возможность прогнозирования ключевых характеристик транспортных потоков с помощью искусственных нейронных сетей (ANN). В данном исследовании выделена практическая значимость полученных результатов как инструмент для выявления и устранения заторов на дорогах.

В рамках исследования [3] рассмотрена имитационная модель, позволяющая воспроизвести максимально близко к реальности транспортные потоки между различными населенными пунктами. Особенность данной работы заключается в оценке экономических факторов развития транспортных сетей, однако масштабировать результаты исследования и внедрить их в существующие системы управления транспортным потоком невозможно.

В работе [4] авторы рассмотрели возможность построения прогнозных моделей для оценки скорости движения транспортного потока на основе неоднородных данных. Особенность предложенного подхода заключается в построении прогноза даже на участках с малым числом наблюдений. При этом, структура транспортной сети, как и другие основные характеристики транспортного потока, не учитывалась при разработке алгоритма.

Авторы статьи [2] исследовали архитектуру графовой свёрточной нейронной сети для прогнозирования параметров транспортных потоков, учитывая дневные и недельные характеристики

транспортного потока и непрерывность дорожной сети. Результаты эксперимента показали высокую точность модели прогнозирования и низкое время работы.

В рамках работы [11] представлена модель прогнозирования транспортного потока на основе рекуррентных нейронных сетей с долговременной и кратковременной памятью на наборе данных с измерениями характеристик транспортного потока на автостраде. Полученные результаты исследования продемонстрировали эффективность модели прогнозирования основных показателей транспортного потока.

Прогнозирование основных параметров транспортного потока представлено в работе [15]. Данный подход использует модель свёрточной нейронной сети для анализа пространственно-временных характеристик транспортного потока. Особенностью данного исследования является возможность оценки вероятности возникновения дорожно-транспортных происшествий в определенный момент времени.

В исследовании [6] продемонстрированы возможности алгоритмов метода k -ближайших соседей (KNN), а также SVM и ANN на наборе данных для прогнозирования дорожных заторов на урбанизированной магистральной дороге. В результате ряда экспериментов показано, что модель ANN эффективнее других моделей справляется с задачей краткосрочного прогнозирования на основе оценки среднеквадратичной ошибки.

Долгосрочная модель прогнозирования основных характеристик транспортного потока предложена в работе [12] и основана на модели функциональной непараметрической регрессии. В структуре модели используется автокорреляционный анализ (ACF) и метод главных компонент для оценки сходства между различными временными рядами характеристик транспортного потока.

В работе [10] представлена модель краткосрочного прогнозирования на основе модели регрессии для индийской системы дорожного движения. Особенность данного исследования заключается в оценке аномального поведения, например, движения в праздничные дни. При этом в качестве индекса стабильности авторы предложили использовать коэффициент вариации и показали, что линейные модели последовательно превосходят непараметрические алгоритмы.

Таким образом, применение прикладных моделей искусственного интеллекта на практике продемонстрировало высокую эффективность в различных задачах прогнозирования основных характеристик транспортного потока. При этом, регрессионные модели позволяют оценить значимость входных признаков, что, соответственно, повышает интерпретируемость модели.

Данная работа направлена на исследование и прогнозирование интенсивности движения транспортного потока на основе набора данных¹, представленного на сайте Kaggle, о наблюдениях за количеством транспортных средств на четырех различных участках перекрестка с интервалом времени, равным одному часу. Прогнозирование интенсивности движения транспортного потока осуществлено посредством реализации многомерной линейной регрессии. Отличительной особенностью данной работы является реализация контроля переобучения с помощью методов регуляризации для построения эффективно масштабируемой прогнозной модели.

Постановка задачи

Объектом исследования является набор данных¹ о наблюдениях за движением на крестообразном перекрестке. Объем набора данных составляет 48120 записей, продолжительность сбора данных с 00:00:00 1 ноября 2015 по 23:00:00 31 июля 2017 года. Доступная информация: уникальный идентификатор записи, номер соединения перекрестка, количество транспортных средств за каждый час, дата и время фиксации наблюдения. Техническое устройство снятия характеристик транспортного потока не указано источником данных. Фрагмент набора данных представлен в таблице 1.

Таблица 1. Фрагмент набора данных

№	Дата-время	Соединение	Транспортные средства	Идентификатор
31891	2016-02-21 19:00:00	3	10	20160221193
15758	2015-12-19 14:00:00	2	5	20151219142
25024	2017-01-08 16:00:00	2	15	201700108162
14780	2015-11-08 20:00:00	2	6	20151108202
45253	2017-03-03 13:00:00	4	9	20170303131

Источник: разработано авторами

Математическая постановка задачи прогнозирования интенсивности движения транспортного потока методами машинного обучения состоит в следующем: пусть существует некоторая зависимость между множеством наблюдений и признаков X и множеством значений зависимой переменной Y .

Необходимо построить модель $a: X \rightarrow Y$ – представляющую собой прогнозную функцию, которая приближает значение $a(x)$ к значениям множества Y на всём обучающем множестве X . При этом, x_1 – идентификационный номер наблюдения; x_2 – номер соединения перекрестка (от 1 до 4); x_3 – дата и вре-

мя фиксации наблюдения в формате год: месяц: число, часы: минуты: секунды; y – количество транспортных средств за каждый час.

Таким образом, получаем задачу краткосрочного прогнозирования, которую будем решать на основе модели многомерной линейной регрессии.

Модель многомерной линейной регрессии

Многомерная линейная регрессия – это модель линейной зависимости одной переменной от нескольких других (независимых) переменных², которая имеет вид:

$$y = \beta_0 + \sum_{j=1}^n \beta_j x_j,$$

где

- x – независимые переменные;
- y – зависимая;
- n – количество признаков;
- β – параметры модели (весовые коэффициенты).

Для построения многомерной линейной регрессии необходимо установить значения коэффициен-

тов, чтобы определить регрессионную модель. При оценке эффективности построенных моделей, как

¹ Набор данных для прогнозирования трафика – Kaggle: Ваш дом для науки о данных, 2021 [Электронный ресурс]. – Режим доступа: <https://www.kaggle.com/fedesoriano/traffic-prediction-dataset> (дата обращения: 04.12.2021).

² Воронцов К. В. Многомерная линейная регрессия, 2019 [Электронный ресурс]. – Режим доступа: http://www.machinelearning.ru/wiki/index.php?title=Многомерная_линейная_регрессия (дата обращения: 05.12.2021).

правило, используются три основных показателя оценки:

- средняя абсолютная ошибка (MAE);
- средняя квадратичная ошибка (MSE);
- корень из средней квадратичной ошибки (RMSE).

Основная проблема параметрических моделей – возможность переобучения параметров, которая выражается в высоком качестве прогноза на данных для обучения, однако при незначительном колебании значений признаков некоторого наблюдения точность прогнозирования может снижаться в несколько раз. В этом случае использовать результаты построенной регрессионной модели для прогнозирования интенсивности движения транспортного потока в другие моменты времени невозможно, т.к. модель теряет способность к масштабированию. В рамках данного исследования для повышения качества модели применены различные подходы к регуляризации.

Регуляризация – это подход, при котором предотвращается переобучение путем решения некорректно поставленной задачи через добавление некоторых дополнительных ограничений к оптимизационному условию.

Для повышения качества модели многомерной линейной регрессии исследуем следующие методы регуляризации:

- 1) регрессия Лассо (L1, LASSO) – взвешивает ошибки параметров по их абсолютному значению и корректирует функцию потерь в соответствующем направлении;
- 2) гребневая регрессия (L2, Ridge) – взвешивает ошибки параметров по их квадрату и регулирует значения коэффициентов переменных функцией потерь в соответствующем направлении;
- 3) регрессия Elastic-Net – метод с двумя регуляризаторами L1 и L2. Данный подход эффективно работает на данных, в которых прогнозное значение сильно зависит от конкретных переменных.

Вычислительные эксперименты

Проведем обучение модели многомерной линейной регрессии на исходных данных, применим рассмотренные выше регуляризаторы и оценим точность по основным метрикам качества. Для адекватной оценки результатов прогнозирования разделим набор данных на обучающий (train) и тестовый (test) наборы в соотношении 4:1 (таблица 2).

Таблица 2. Метрики качества прогноза на исходных данных

Метрика	Многомерная линейная регрессия		Методы регуляризации					
			LASSO		Ridge		Elastic Net	
	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>
MAE	9,9887	9,8261	9,8154	9,6761	10,0268	9,8474	10,7423	10,7443
MSE	182,4048	178,2592	188,8495	185,6962	182,5926	178,0251	240,8773	241,5413
RMSE	13,5057	13,3514	13,7423	13,6270	13,5127	13,3426	15,5202	15,5416

Источник: разработано авторами

Результаты обучения прогнозных моделей показали, что наименьшую среднеквадратичную ошибку на тестовых данных показала модель Ridge: на 0,07% точность выше по сравнению с исходной моделью многомерной регрессии, на 2,09% точность выше модели с регуляризацией LASSO, а также на 16,45% лучше модели с ElasticNet. Таким образом,

применение регуляризации для контроля переобучения позволило повысить точность прогнозирования.

Исследуем влияние масштаба признаков на качество прогнозирования и приведем их к единому диапазону. Применим стандартизацию данных и на них проведем обучение модели многомерной линейной регрессии с учетом регуляризации (таблица 3).

Таблица 3. Метрики качества прогноза после нормализации данных

Метрика	Многомерная линейная регрессия		Методы регуляризации					
			LASSO		Ridge		ElasticNet	
	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>	<i>train</i>	<i>test</i>
MAE	0,4762	0,4701	0,7403	0,7471	0,4767	0,4705	0,6993	0,7051
MSE	0,4164	0,4070	0,9965	1,0141	0,4164	0,4069	0,9097	0,9242
RMSE	0,6453	0,6380	0,9982	1,0070	0,6453	0,6380	0,9538	0,9614

Источник: разработано авторами

После нормализации данных наименьшую среднеквадратичную ошибку на тестовых данных показали модель многомерной линейной регрессии совместно с регуляризацией Ridge: на 57,84% лучше, чем метод регуляризации LASSO и на 50,69% лучше, чем метод регуляризации ElasticNet.

Среднеквадратичная ошибка модели многомерной линейной регрессии с регуляризацией Ridge после применения нормализации данных уменьшилась на 95,22%. Таким образом, предварительная обработка данных в виде стандартизации признаков позволила повысить качество прогнозной модели.

Исследуем связь между остатками модели (отклонение прогнозных значений от истинных) и предсказанными значениями для оценки качества модели. На рисунке 1 представлены графики остат-

ков на исходных данных, а на рисунке 2 – остатки на нормализованных признаках. Полученные результаты позволяют отследить нелинейность и выбросы, проверить случайность распределения ошибки.

Графики, представленные на рисунках 1 и 2, показали высокую плотность точек, близких к нулю, и низкую плотность точек, удаленных от нуля. Кроме того, область определения предсказанных значений на тестовых данных входит в область определения на обучающих данных, что свидетельствует о высоком качестве построенных прогнозных моделей. Рисунок 2b и 2c показывает, что метод регуляризации LASSO сократил все коэффициенты до 0, что является следствием минимизации ошибки перекрестной проверки при большой усадке.

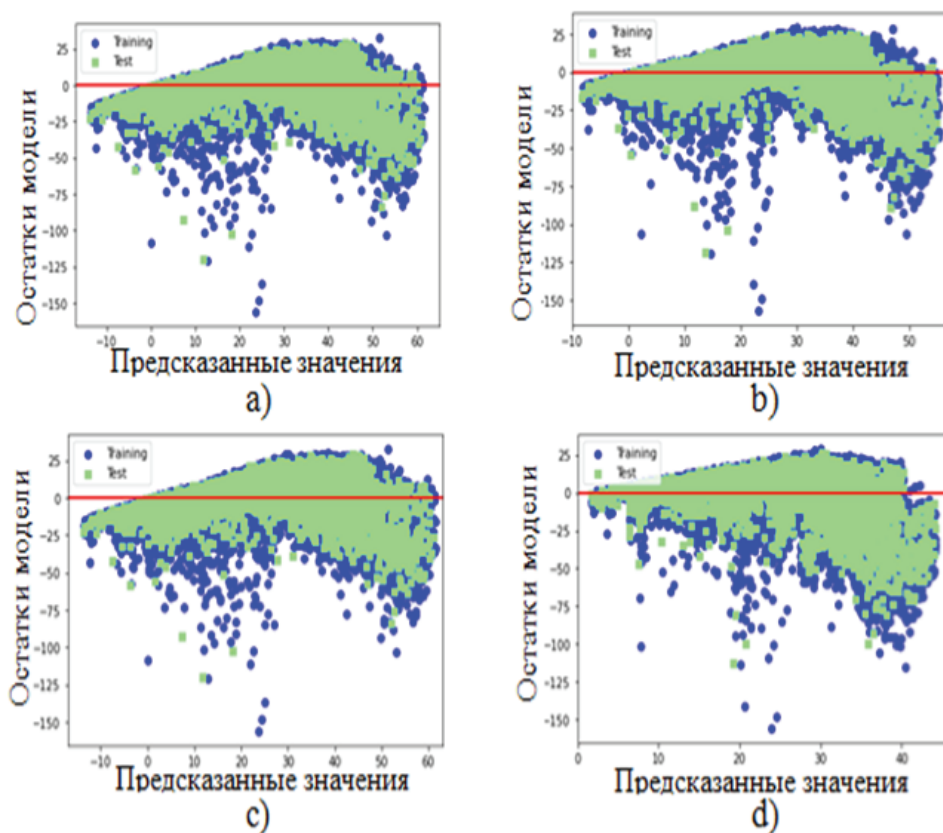


Рисунок 1. Взаимосвязь ошибки прогноза с величиной предсказанных значений (a – Линейная регрессия, b – LASSO, c – Ridge, d – ElasticNet)

Источник: разработано авторами

Таким образом, на исследуемом наборе данных для прогнозирования интенсивности движения транспортного потока эффективнее работает многомерная линейная регрессия после нормализации данных, а лучшим методом контроля переобучения является Ridge-регуляризация.

Методика применения разработанной модели прогнозирования интенсивности движения транспортного потока на новом перекрестке следующая:

1. Для крестообразного перекрестка нумеруем соединения от 1 до 4, далее фиксируем не менее 45 тыс. наблюдений о количестве транспортных средств за каждый час с сохранением информации о дате и времени.

2. Согласно представленной модели многомерной регрессии с Ridge-регуляризацией обновляем значения параметров $\beta_j, j = 1, \dots, n$ на основе метода наименьших квадратов.

3. Применяем прогнозную модель: передаем значение новой даты и времени для вычисления

прогноза интенсивности движения транспортного потока в соответствующий момент времени.

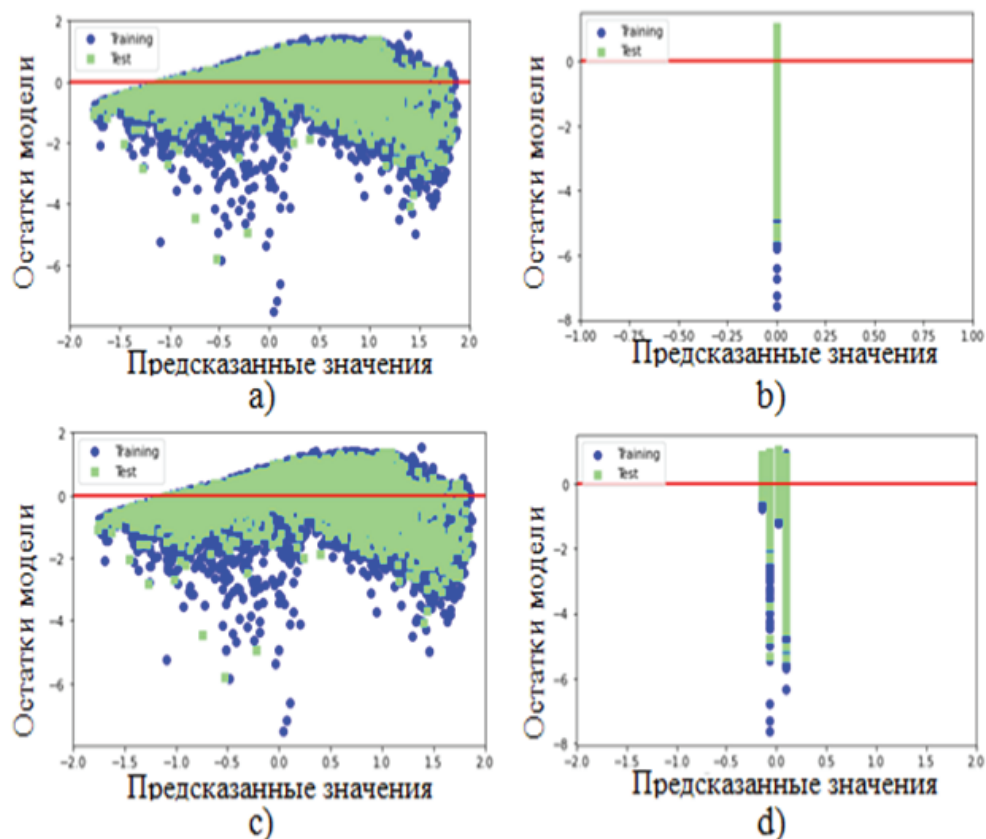


Рисунок 2. Взаимосвязь ошибки прогноза с величиной предсказанных значений после нормализации переменных (a – Линейная регрессия, b – LASSO, c – Ridge, d – ElasticNet)

Источник: разработано авторами

Заключение

По результатам прогнозирования интенсивности движения транспортного потока посредством реализации многомерной линейной регрессии и контроля переобучения с помощью методов регуляризации можно сделать следующие выводы:

- наименьшую ошибку на тестовых исходных данных показала модель с Ridge-регуляризацией: на 0,07% точность выше по сравнению с исходной моделью многомерной регрессии;
- предварительная обработка данных в виде стандартизации признаков позволила повысить качество прогнозной модели на 57,84%;
- построенная модель многомерной линейной регрессии позволяет спрогнозировать интенсивность движения транспортного потока на перекрестке со средней квадратичной ошибкой, равной 0,638.

Литература

1. Агафонов А. А., Мясников В. В. Оценка и прогнозирование параметров транспортных потоков с использованием композиции методов машинного обучения и моделей прогнозирования временных рядов // Компьютерная оптика. – 2014. – том 38. – № 3. – С. 539–549, <https://doi.org/10.18287/0134-2452-2014-38-3-539-549>. EDN: TGFKRR.

2. Агафонов А. А., Юмаганова А. С. Прогнозирование транспортного потока с использованием графовых нейронных сетей // Информационные технологии и нанотехнологии (ИТНТ-2020): сборник трудов по материалам VI Международной конференции и молодежной школы, Самара, 26–29 мая 2020 г.: в 4-х т. – Самара. – 2020. – Т. 4. – Науки о данных. – С. 143–152. EDN: JYJJMA.
3. Владимирская И. П. Прогнозирование транспортных потоков в рыночной экономике // Научный вестник Московского государственного технического университета гражданской авиации. – 2009. – № 147. – С. 161–165. EDN: LDHNRX.
4. Пупырев С., Пронченков А. Прогнозирования загруженности автомобильных дорог // IV Российская летняя школа по информационному поиску (RuSSIR'2010): труды Четвертой Российской конференции молодых ученых по информационному поиску, Воронеж, 13–18 сентября 2010 г. – Воронеж. – 2010. – С. 64–78.
5. Эшмурадов Д. Э., Мухамеджанов А. А. Анализ и прогнозирование транспортных потоков с применением информационных технологий // Теория и практика современной науки. – 2021. – № 12. – С. 424–435, https://doi.org/10.46566/2412-9682_2021_78_424. EDN: SNSDXV.
6. Bartlett Z. et al. (2018) A Machine Learning Based Approach for the Prediction of Road Traffic Flow on Urbanised Arterial Roads. *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/ DSS)*. Pp. 1285–1292, <https://doi.org/10.1109/HPCC/SmartCity/DSS.2018.00215>. (In Eng.).
7. Castro-Neto M. et al. (2009) Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications*. Vol. 36. No. 3, pp. 6164–6173, <https://doi.org/10.1016/j.eswa.2008.07.069>. (In Eng.).
8. Chan K. Y. et al. (2012) Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm, *IEEE Transactions on Intelligent Transportation Systems*. Vol. 13. No. 2, pp. 644–654, <https://doi.org/10.1109/TITS.2011.2174051>. (In Eng.).
9. Lippi M., Bertini M., Frasconi P. (2016) Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning, *IEEE Transactions on Intelligent Transportation Systems*. Vol. 14, No. 2, pp. 871–882, <https://doi.org/10.1109/TITS.2013.2247040>. (In Eng.).
10. Rajendran S., Ayyasamy B. (2020) Short-term traffic prediction model for urban transportation using structure pattern and regression: an Indian context. *SN Applied Sciences*. Vol. 2, 1159, <https://doi.org/10.1007/s42452-020-2946-2>. (In Eng.).
11. Rezaa S. et al. (2022) A Multi-head Attention-based Transformer Model for Traffic Flow Forecasting with a Comparative Analysis to Recurrent Neural Networks. *Expert Systems with Applications*. Vol. 202, 117275, <https://doi.org/10.1016/j.eswa.2022.117275>. (In Eng.).
12. Su F. et al. (2016) Long-term forecasting oriented to urban expressway traffic situation. *Advances in Mechanical Engineering*. Vol. 8. No. 1, pp. 1–16, <https://doi.org/10.1177/1687814016628397>. (In Eng.).
13. Vlahogianni E. I., Karlaftis M. G., Golias J. C. (2014) Short-term traffic forecasting: where we are and where we're going. *Transportation Research Part C: Emerging Technologies*. Vol. 43, pp. 3–19, <https://doi.org/10.1016/j.trc.2014.01.005> (In Eng.).
14. Williams B. M., Hoel L. A. (2003) Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *Journal of Transportation Engineering*. Vol. 129, No. 6, pp. 664–672, [https://doi.org/10.1061/\(ASCE\)0733-947X\(2003\)129:6\(664\)](https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664)). (In Eng.).
15. Yafang L. et al. (2022). A Grey Convolutional Neural Network Model for Traffic Flow Prediction under Traffic Accidents. *Neurocomputing*. Vol. 500. No. 1, pp. 761–775, <https://doi.org/10.1016/j.neucom.2022.05.072>. (In Eng.).

References

1. Agafonov, A. A., Myasnikov, V. V. (2014) [Estimation and forecasting of traffic flow parameters using the composition of machine learning methods and time series forecasting models]. *Komp'yuternaya optika* [Computer Optics]. Vol. 38. No. 3, pp. 539–549, <https://doi.org/10.18287/0134-2452-2014-38-3-539-549>. (In Russ.).
2. Agafonov, A. A., Yumaganova, A. S. (2020) [Prediction of traffic flow using graph neural networks]. *Informatsionnyye tekhnologii i nanotekhnologii* [Information technology and nanotechnology]. Vol. 4, pp. 143–152. (In Russ.).
3. Vladimirskaya, I. P. (2009) [Forecasting traffic flows in a market economy]. *Nauchnyy vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta grazhdanskoy aviatsii* [Scientific Bulletin of the Moscow State Technical University of Civil Aviation]. Vol. 147, pp. 161–165. (In Russ.).
4. Pupyrev, S., Pronchenkov, A. (2010) [Forecasting traffic congestion]. *Trudy konferentsii molodykh*

uchennykh RUSSIR 2010 [Proceedings of the Conference of Young scientists RUSSIR 2010]. Voronezh, pp. 64–78. (In Russ.).

5. Eshmuradov, D. E., Muxamedjanov, A. A. (2021) [Analysis and forecasting of traffic flows using information technologies]. *Teoriya i praktika sovremennoy nauki* [Theory and practice of modern science]. Vol. 12, pp. 78–83. (In Russ.).

6. Bartlett, Z. et al. (2018) A Machine Learning Based Approach for the Prediction of Road Traffic Flow on Urbanised Arterial Roads. *2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/ DSS)*, pp. 1285–1292, <https://doi.org/10.1109/HPCC/SmartCity/DSS.2018.00215>. (In Eng.).

7. Castro-Neto, M. et al. (2009) Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. *Expert Systems with Applications*. Vol. 36. No. 3, pp. 6164–6173, <https://doi.org/10.1016/j.eswa.2008.07.069>. (In Eng.).

8. Chan, K. Y. et al. (2012) Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm, *IEEE Transactions on Intelligent Transportation Systems*. Vol. 13. No. 2, pp. 644–654, <https://doi.org/10.1109/TITS.2011.2174051>. (In Eng.).

9. Lippi, M., Bertini, M., Frasconi, P. (2016) Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning, *IEEE Trans. Intelligent Transportation Systems*. Vol. 14, No. 2, pp. 871882, <https://doi.org/10.1109/TITS.2013.2247040>. (In Eng.).

10. Rajendran, S., Ayyasamy, B. (2020) Short-term traffic prediction model for urban transportation using structure pattern and regression: an Indian context. *SN Applied Sciences*. Vol. 2, pp. 1–11, <https://doi.org/10.1007/s42452-020-2946-2>. (In Eng.).

11. Rezaa, S. et al. (2022) A Multi-head Attention-based Transformer Model for Traffic Flow Forecasting with a Comparative Analysis to Recurrent Neural Networks. *Expert Systems with Applications*. Vol. 202, pp. 1–11. 117275, <https://doi.org/10.1016/j.eswa.2022.117275>. (In Eng.).

12. Su, F. et al. (2016) Long-term forecasting oriented to urban expressway traffic situation. *Advances in Mechanical Engineering*. Vol. 8. No. 1, pp. 1–16, <https://doi.org/10.1177/1687814016628397>. (In Eng.).

13. Vlahogianni, E. I., Karlaftis, M. G., Golias, J. C. (2014) Short-term traffic forecasting: where we are and where we're going. *Transportation Research Part C: Emerging Technologies*. Vol. 43, pp. 3–19, <https://doi.org/10.1016/j.trc.2014.01.005> (In Eng.).

14. Williams, B. M., Hoel, L. A. (2003) Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results. *Journal of Transportation Engineering*. Vol. 129, No. 6, pp. 664672, [https://doi.org/10.1061/\(ASCE\)0733-947X\(2003\)129:6\(664\)](https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664)). (In Eng.).

15. Yafang, L. et al. (2022) A Grey Convolutional Neural Network Model for Traffic Flow Prediction under Traffic Accidents. *Neurocomputing*. Vol. 500. No. 1, pp. 761–775, <https://doi.org/10.1016/j.neucom.2022.05.072>. (In Eng.).

Информация об авторах:

Ирина Павловна Болодурина, доктор технических наук, профессор, заведующий кафедрой прикладной математики, Оренбургский государственный университет, Оренбург, Россия

ORCID ID: 0000-0003-0096-2587, **Researcher ID:** E-7202-2015, **Scopus ID:** 55809754500, **РИНЦ (AuthorId):** 118837

e-mail: ipbolodurina@yandex.ru

Лариса Михайлова Анциферова, кандидат педагогических наук, доцент кафедры прикладной математики, Оренбургский государственный университет, Оренбург, Россия

ORCID ID: 0000-0003-0181-7402, **Scopus ID:** 57202401235, **РИНЦ (AuthorId):** 635081

e-mail: antsiferova_68@mail.ru

Любовь Сергеевна Гришина, старший преподаватель кафедры прикладной математики, Оренбургский государственный университет, Оренбург, Россия

ORCID ID: 0000-0003-2752-7198, **Scopus ID:** 57205597695, **РИНЦ (AuthorId):** 1084077

e-mail: grishina_ls@inbox.ru

Вклад соавторов:

Болодурина И. П. – общее научное руководство исследованием, включая постановку цели и задачи, консультирование по методике проведения исследования, формулировка выводов.

Анциферова Л. М. – обзор метода машинного обучения, применения в методе регрессоров, анализ и обсуждение результатов.

Гришина Л. С. – выбор и анализ набора данных, программная реализация, анализ и обработка экспериментальных вычислений.

Статья поступила в редакцию: 01.06.2022; принята в печать: 23.09.2022.

Авторы прочитали и одобрили окончательный вариант рукописи.

Information about the authors:

Irina Pavlovna Bolodurina, Doctor of Technical Sciences, Professor, Head of the Department of Applied Mathematics, Orenburg State University, Orenburg, Russia

ORCID ID: 0000-0003-0096-2587, **Researcher ID:** E-7202-2015, **Scopus ID:** 55809754500, **РИИЦ (AuthorId):** 118837

e-mail: ipbolodurina@yandex.ru

Larisa Mikhailovna Antsiferova, Candidate of Pedagogical Sciences, Associate Professor of the Department of Applied Mathematics, Orenburg State University, Orenburg, Russia

ORCID ID: 0000-0003-0181-7402, **Scopus ID:** 57202401235, **РИИЦ (AuthorId):** 635081

e-mail: antsiferova_68@mail.ru

Lyubov Sergeevna Grishina, Lecturer of the Department of Applied Mathematics, Orenburg State University, Orenburg, Russia

ORCID ID: 0000-0003-2752-7198, **Scopus ID:** 57205597695, **РИИЦ (AuthorId):** 1084077

e-mail: grishina_ls@inbox.ru

Contribution of the authors:

Bolodurina I. P. – general scientific management of the research, including setting goals and objectives, consulting on the methodology of the research, formulation of conclusions.

Antsiferova L. M. – review of machine learning methods, applications in the regressor method, analysis and discussion of the results.

Grishina L. S. – selection and analysis of a data set, software implementation, analysis and processing of experimental calculations.

The paper was submitted: 01.06.2022.

Accepted for publication: 23.09.2022.

The authors have read and approved the final manuscript.