

УДК 004.891.2

ПРИМЕНЕНИЕ МЕТОДА ЛИНЕЙНОЙ РЕГРЕССИИ ДЛЯ РЕШЕНИЯ ТРАНСПОРТНЫХ ЗАДАЧ

Невейков А. С.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

Научный руководитель: Ролч О. Ч. - канд. техн. наук, доцент, доцент кафедры ПИКС

Аннотация. Исследован способ решения транспортных задач методом линейной регрессии с применением машинного обучения. Создана модель, предсказывающая интенсивность трафика на дорожном участке. Оценено влияние исследуемых параметров на загруженность участка дороги, рассчитана погрешность данной оценки.

Ключевые слова: линейная регрессия, машинное обучение, транспортная задача, линейный тренд

Введение. При разработке городских инфраструктурных проектов возникает необходимость решать такие транспортные задачи, как определение числа пешеходов или водителей в зависимости от погодных условий, времени года, дня недели. Данная работа предназначена для анализа применимости метода линейной регрессии с использованием машинного обучения в решении подобного класса задач.

Основная часть. Проблема применения рассматриваемого метода – это нелинейность данных. Транспортные потоки цикличны, причем цикличность одновременно может проявляться внутри различных временных периодов (сутки, неделя, год) и из-за различных иных факторов. В час-пик на дорогах больше автомобилей, летом в парках пешеходов больше, чем зимой и т.д.

Линейная регрессия представляет собой либо линейную аппроксимацию вида:

$$y = ax + b, \quad (1)$$

где a – угловой коэффициент,

b – точка пересечения с осью y [1],

либо многомерную линейную модель следующего вида:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots, \quad (2)$$

что явно не подходит для поставленной задачи.

Во-первых, для применения метода линейной регрессии выполнено преобразование данных в соответствии с новыми базисными функциями; взята многомерная модель вида:

$$y = a_0 + a_1x_1 + a_2x_2^2 + a_3x_3^3 \quad (3)$$

и построены x_1, x_2, x_3 и т.д. на основе имеющегося одномерного входного значения x .

То есть $x_n = f_n(x)$, где $f_n(x)$ – некая функция, выполняющая преобразование данных. Тогда, при $f_n(x) = x^n$, модель превращается в полиномиальную регрессию и остается линейной, так как линейность относится к тому, что коэффициенты a_n не умножаются и не делятся друг на друга [4].

Смысл преобразования в том, что для одномерных значений x построена проекция на многомерное пространство, и с помощью линейной аппроксимации стало возможным отображать сложные зависимости между x и y .

Во-вторых, разработана и обучена на реальных данных [5, 6] модель, которая, анализируя информацию о погоде, времени года и другие факторы, предсказывает интенсивность трафика на велосипедной дорожке (рисунок 1). Оценен вклад исследуемых параметров в количество велосипедистов и рассчитана погрешность каждого из них (рисунок 2).

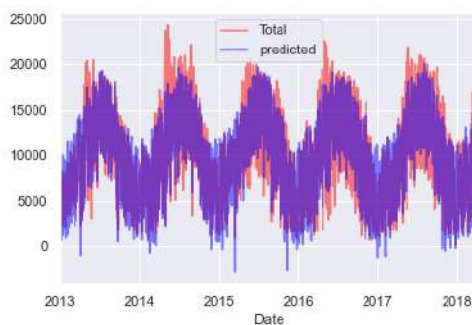


Рисунок 1 – Результат предсказания модели.

Красный цвет – реальное число велосипедистов, синий цвет предсказание модели, фиолетовый цвет пересечение реальных данных и предсказания

	effect	error
Mon	-2155.0	273.0
Tue	-1761.0	283.0
Wed	-1782.0	287.0
Thu	-2377.0	279.0
Fri	-3677.0	276.0
Sat	-9034.0	271.0
Sun	-9348.0	266.0
holiday	-4843.0	475.0
daylight_hrs	429.0	28.0
PRCP	-28136.0	1972.0
dry day	2010.0	108.0
Temp (c)	1615.0	64.0
annual	122.0	28.0

Рисунок 2 – Оценка исследуемых параметров.

Левая колонка - список параметров, средняя колонка - численная оценка влияния параметра
правая колонка - рассчитанная погрешность

Заключение. Метод линейной регрессии с использованием машинного обучения применим для решения транспортных задач. Несмотря на отсутствие нелинейных трендов в пределах каждой из переменных, которые не могут быть учтены в этой модели, и эффекты от совместного влияния нескольких параметров, обнаружены устойчивые тенденции относительно изучаемых параметров (с учетом погрешности), заключающиеся в сокращении трафика в выходные и праздничные дни, и при осадках; увеличении трафика в будние дни, особенно в начале недели, в теплые дни, при отсутствии осадков и увеличении длины светового дня. Также обнаружена ежегодная цикличность транспортного потока, что говорит о применимости метода линейной регрессии для решения практических транспортных задач.

Список литературы

1. Портал знаний [Электронный ресурс]. – Режим доступа <http://statistica.ru/theory/osnovy-lineynoy-regressii/>
2. Proglib [Электронный ресурс]. – Режим доступа <https://proglib.io/p/linear-regression>
3. J. VanderPlas, Python для сложных задач: наука о данных и машинное обучение. – СПб.: Питер, 2018. – 576 с
4. PythonRu [Электронный ресурс]. – Режим доступа <https://pythonru.com/uroki/linear-regression-sklearn>
5. Sesttle [Электронный ресурс]. – Режим доступа <https://data.seattle.gov/Transportation/Fremont-Bridge-Bicycle-Counter/65db-xm6k>
6. ClimateDataOnline [Электронный ресурс]. – Режим доступа <https://www.ncdc.noaa.gov/cdo-web>

UDC 004.891.2

APPLICATION OF THE LINEAR REGRESSION METHOD FOR SOLVING TRANSPORT PROBLEMS

Neveikov A.S.

Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Rolich O. Ch. – Ph.D. assistant professor, associate professor of the department of ICSD

Annotation. Research method for solving problems by linear regression using machine learning. A model has been created that predicts the movement of vehicles on the road section. The influence of the dependencies of the parameters on the congestion of the road section, the expected error of this assessment, is estimated.

Keywords: linear regression, machine learning, transportation problem, linear trend