**Применение метода линейной регрессии**

**Для решения транспортных задач**

студент гр. 014301 Невейков А. С.,

*Научный руководитель - канд. техн. наук Ролич О. Ч.*

Белорусский государственный университет информатики и радиоэлектроники

Минск, Беларусь

При разработке инфраструктурных проектов, возникает необходимость решать транспортные задачи, как например: определение числа пешеходов или водителей в зависимости от погодных условий, времени года, дня недели и других. Данная работа предназначена для анализа применимости метода линейной регрессии с применением машинного обучения для решения подобного класса задач.

Основная проблема применения рассматриваемого метода – это нелинейность данных: транспортные потоки цикличны, причем цикличность одновременно может проявляться внутри различных временных периодов (сутки, неделя, года) и из-за различных факторов. В час-пик на дорогах будет больше автомобилей, тогда как в парках летом пешеходов будет больше чем зимой. Линейная регрессия же представляет собой либо линейную аппроксимацию вида , где – угловой коэффициент, – точка пересечения с осью , либо многомерную линейную модель следующего вида: … , 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 … и построены , и т.д. на основе имеющегося одномерного входного значения . То есть , где – некая функция, выполняющая преобразование данных. Тогда, при , модель превращается в полиномиальную регрессию и остается линейной, так как линейность относится к тому, что коэффициенты не умножаются и не делятся друг на друга. Смысл преобразования в том, что для одномерных значений построена проекция на многомерное пространство, так что с помощью линейной аппроксимации стало возможным отображать сложные зависимости между и .

Во-вторых, разработана модель, которая на основании данных о погоде, времени года и других факторах предсказывает количество трафика на велосипедной дорожке. И оценено какой вклад вносит каждый из параметров в количество велосипедистов (рис. 1).

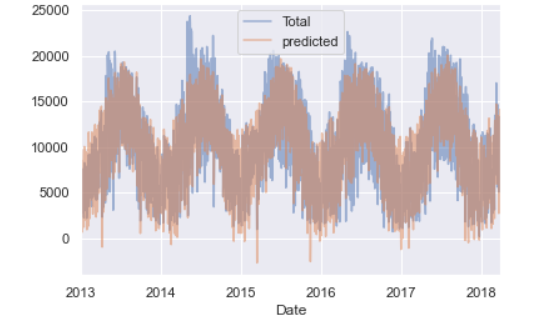


Рисунок 1 – Результат предсказания модели

Таким образом можно говорить о применимости метода линейной регрессии для решения транспортных задач. Не смотря на отсутствие нелинейных трендов в пределах каждой из переменных, которые не могут быть учтены в этой модели, и эффекты от совместного влияния нескольких параметров, были обнаружены устойчивые тенденции, относительно изучаемых параметров (с учетом погрешности), что говорит о применимости данного метода.

*Литература*

1. Sesttle [Электронный ресурс]. – Режим доступа https://data.seattle.gov/Transportation/Fremont-Bridge-Bicycle-Counter/65db-xm6k
2. Climate Data Online [Электронный ресурс]. – Режим доступа https://www.ncdc.noaa.gov/cdo-web
3. Python Data Science Handbook: Essential Tools for Working with Data by J. Vander Plas