

Санкт-Петербургский государственный электротехнический университет им.

В.И. Ульянова (Ленина)

ПРОГНОЗИРОВАНИЕ СКОРОСТИ ДВИЖЕНИЯ АВТОМОБИЛЬНОГО ПОТОКА С ИСПОЛЬЗОВАНИЕМ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ДАННЫХ

Выполнила:

Алясова Анастасия Николаевна, гр. 7381

Руководитель:

Геппенер Владимир Владимирович, д.т.н., профессор каф. МОЭВМ

Консультант:

Борисенко Константин Алексеевич. , к.т.н., ст. преподаватель каф. МОЭВМ

Санкт-Петербург, 2021

Актуальность работы

Прогнозирование скорости дорожного потока является одним из способов устранения избыточного дорожного трафика, так как позволяет заранее водителям вычислять более быстрые по времени маршруты.

Может положительно повлиять:

- ☺ на решение проблемы перегруженности автомобильного движения
- ☺ на темпы экономического развития обслуживаемых дорогами территорий

Цели и задачи

Цель работы.

При помощи метода интеллектуального анализа данных разработать метод для прогнозирования скорости движения автомобильного потока для устранения избыточного дорожного трафика и улучшения ситуаций на дорогах.

Задачи.

- 1) Исследовать выбранную дорожную сеть для выявления основных тенденции и закономерностей автомобильного движения.
- 2) Построить базу данных об автомобильном трафике, основанную на данных о средней скорости машин, днях недели, времени.
- 3) Рассмотреть некоторые алгоритмы прогнозирования.
- 4) Построить модели для прогнозирования скорости трафика с помощью ИАД
- 5) Выбрать метрики для оценки модели и провести тестирование.

Исследование дорожной сети для выявления тенденции и закономерностей

Рассмотрим статистику по дням недели в промежуток времени с 4:00 до 5:00.

	ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
Минимальная скорость	32	32	34	29	40	32	34
Максимальная скорость	54	60	62	55	56	54	56
Средняя скорость	44	48	46	46	47	44	44
Медиана	44	50	46	46	47	43	43
Стандартное отклонение	5	7	7	6	5	5	6

Рассмотрим статистику по дням недели в промежуток времени с 18:00 до 19:00.

	ПН	ВТ	СР	ЧТ	ПТ	СБ	ВС
Минимальная скорость	19	16	20	22	24	12	26
Максимальная скорость	38	43	42	37	35	40	40
Средняя скорость	31	31	30	30	29	30	34
Медиана	31	33	30	31	29	32	34
Стандартное отклонение	4	6	5	3	3	7	3

Построение базы данных

База данных

traffic_speed	day	week_day	time
41.348069	0	saturday	0.00
37.241356	0	saturday	0.15
33.832750	0	saturday	0.30
40.403444	0	saturday	0.45
49.843057	0	saturday	1.00
...
46.160678	60	wednesday	22.45
48.828092	60	wednesday	23.00
45.487081	60	wednesday	23.15
39.848074	60	wednesday	23.30
42.930817	60	wednesday	23.45

Входные данные

day	week_day	time
0	2	0
0	2	1
0	2	2
0	2	3
0	2	4
...
60	6	91
60	6	92
60	6	93
60	6	94
60	6	95

Целевые данные

41.348069
37.241356
33.832750
40.403444
49.843057
...
46.160678
48.828092
45.487081
39.848074
42.930817

Выбор алгоритма прогнозирования

Были рассмотрены алгоритмы прогнозирования:

- Линейная регрессия
- Метод k-ближайших соседей
- Случайный лес
- Деревья решений
- Искусственная нейронная сеть

Выбор алгоритма прогнозирования

Обучение	Значение погрешности, км/ч	LR	K-NN	RFR	DTR	NNS
Обучение на всех данных		58 %	85 %	85 %	77 %	60 %
Обучение на всех данных с 18 - 19		80 %	82 %	88 %	86 %	61 %
Обучение по всем данным определенного дня недели		58 %	86 %	88 %	75 %	53 %

Построение модели прогнозирования

Ансамбль моделей случайного леса и метода k-ближайших соседей

1) Метод k-ближайших соседей

```
KNeighborsRegressor(n_neighbors=4, algorithm='ball_tree',  
                    weights='distance', p=1)
```

2) Случайный лес

```
RandomForestRegressor(n_estimators=500, min_weight_fraction_leaf=0,  
                      max_features='sqrt', random_state=0)
```


Тестирование и оценка модели

Результаты тестирования модели при различной погрешности.

Обучение \ Значение погрешности, км/ч	5	4	3
Обучение на всех данных	86 %	78 %	67 %
Обучение на всех данных с 18 - 19	91 %	84 %	61 %
Обучение по всем данным опред. дней недели	91 %	83 %	66 %

Тестирование и оценка модели

Результаты предсказания скорости с 18:00-19:00 для последнего дня.

Обучение	Значение погрешности, км/ч	5	4	3
Обучение на всех данных		100 %	75 %	75 %
Обучение на данных дней недели		0 %	0 %	0 %
Обучение на данные за неделю до дня предсказания		75 %	75 %	75 %
1 способ		75 %	75 %	75 %
Обучение на временном интервале на всех данных		50 %	50 %	50 %
Обучение на временном интервале на данных дней недели		0 %	0 %	0 %
Обучение на временном интервале на данных недели до дня предсказания		50 %	50 %	50 %
2 способ		50 %	50 %	50 %
3 способ		100 %	100 %	75 %

Апробация

Исходный код выложен в открытый доступ в репозитории проекта:

<https://github.com/Aliaska-An/Road-prediction>

Данные тестирования находятся в открытом доступе:

<https://github.com/Aliaska-An/Road-prediction>

Заключение

- Определены и подробно рассмотрены необходимые для решения задачи математические модели и алгоритмы
- Создано собственное решение по прогнозированию скорости автомобильного потока
- Экспериментальное исследование скорости работы программы показало, что программа работает быстро и корректно
- Проведён анализ результатов работы программы
 - Наилучший результат для предсказания обучается на всех данных, данных за неделю до дня предсказания и на данных определенного интервала времени дня недели