

ОБЗОР ПОДХОДОВ К ПРИМЕНЕНИЮ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ВЫЯВЛЕНИЯ ФАКТОРОВ, ВЛИЯЮЩИХ НА РИСК ВОЗНИКНОВЕНИЯ ДТП

© 2018 Д. С. Донченко, Н. П. Садовникова, Д. С. Парыгин

Волгоградский государственный технический университет (г. Волгоград, Россия)

В работе рассматриваются существующие подходы к применению алгоритмов искусственного интеллекта для анализа данных о дорожно-транспортных происшествиях в разных странах мира. Предложен подход к использованию искусственного интеллекта для анализа данных о ДТП в России с целью выявления факторов, увеличивающих риск возникновения ДТП.

Ключевые слова: машинное обучение, искусственный интеллект, анализ данных.

Согласно официальной статистике, предоставленной дорожной полицией России на сайте <http://stat.gibdd.ru/> [1], число дорожно-транспортных происшествий уменьшается почти каждый год в течение последних 12 лет. По данным ВОЗ [2], смертность от дорожно-транспортных происшествий в 2015 году составила 15.8 на 100 тысяч населения. Для сравнения, в 2005 году этот показатель составлял 23,6 на 100 тысяч человек [3]. Однако, число смертей на дорогах остается очень высоким по сравнению со многими развитыми странами – для сравнения, в Швеции и Великобритании показатели смертности не превышают 3 человек на 100 тысяч населения, в среднем по странам ЕС этот показатель не превышает 10 человек на 100 тысяч населения. Кроме того, дорожная полиция готовится к резким дорожным происшествиям в России после 2018 года. Основными причинами этого являются старение водителей и парка автомобилей, неконтролируемая моторизация населения на фоне проблем в экономике и ряде других факторов (4).

Концепция использования моделей прогнозирования дорожно-транспортных происшествий достаточно широко распространена и имеет множество успешных применений в управлении безопасностью дорожного движения. В США Tennessee Highway Patrol (ТНП) применил разработанное IBM решение для прогнозирующего моделирования, которое привело к снижению дорожно-транспортных происшествий на 6 процентов

и увеличению арестов DUI (driving under influence – вождение под влиянием алкоголя или наркотиков) на 34 процента [5]. Эта модель использует исторические данные о дорожно-транспортных происшествиях, статистику DUI-арестов, статистику погоды и данные о популярных событиях для прогнозирования вероятности дорожно-транспортных происшествий в определенных местах.

Финляндия использует систему под названием TARVA для оценки факторов дорожной инфраструктуры, влияющих на безопасность дорожного движения [6]. TARVA имеет более чем 80-процентную точность в прогнозировании данных об авариях за три года и использует эмпирический байесовский метод для оценки эффектов безопасности дорожного движения.

В статье A review of accident prediction models for road intersections [7] описаны преимущества различных типов моделей прогнозирования аварий, таких как модели линейной и логистической регрессии, модели Пуассона, методика деревьев классификации и регрессии (CART), модели с отрицательным биномиальным индексом и модели случайных эффектов.

Однако, нам не удалось найти статьи, описывающие подобные предсказательные модели, которые показывают хороший результат в прогнозировании дорожно-транспортных происшествий в России. Авторы исследования «Анализ причин и последствий дорожно-транспортных происшествий» предлагают подход для выявления факторов, существенно влияющих на риск дорожно-транспортного происшествия при решении проблемы повышения безопасности дорожного движения [8]. В работе «Повышение эффективности прогнозирования ДТП на автомобильных дорогах вне населенных пунктов на основе разработки экспертной

Донченко Дмитрий Сергеевич – Волгоградский государственный технический университет, аспирант, dmitrydonchenko92@gmail.com.

Садовникова Наталья Петровна – Волгоградский государственный технический университет, д. т. н., профессор каф. САПРиПК.

Парыгин Данила Сергеевич – Волгоградский государственный технический университет, к. т. н., доцент каф. САПРиПК.

системы» [9] авторы проверили некоторые распространенные методы прогнозирования аварий на реальных данных с дорог Алтайского края за 2011-2014 годы. Качество прогноза было низким из-за отсутствия данных о дорожных условиях. В то же время, некоторые решения для прогнозирования дорожно-транспортных происшествий находятся в стадии разработки, но детали их реализации еще не обнародованы ([10], [11]).

В данной работе, мы воспользовались открытыми данными о ДТП, которые доступны на официальном сайте ГИБДД [1]. Для анализа использовались данные о ДТП, произошедших в России с начала 2015 года по апрель 2018 года. Скачанные данные для удобства работы с ними были занесены в базу данных. Для визуализации данных была использована технология Apache Superset [12].

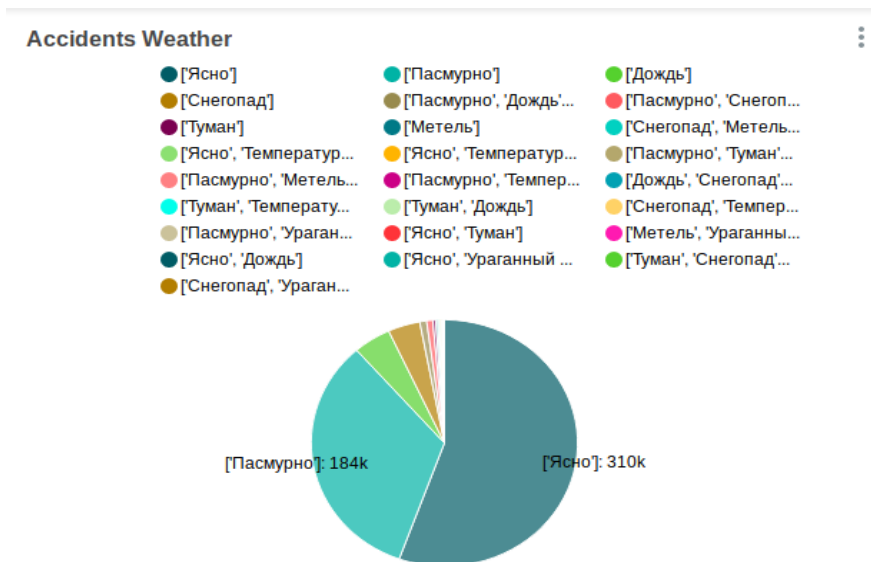


Рисунок 1. Погодные условия, при которых возникали ДТП в России в 2015-2018 гг.



Рисунок 2. Типы ДТП в России за 2015-2018 гг.

Мы выделили следующие показатели из имеющихся данных о ДТП в России, которые могут быть использованы в модели для прогнозирования риска ДТП:

1. Погодные условия:
 - a. Без осадков
 - b. Дождь
 - c. Снег
 - d. Метель
 - e. Туман

2. Время суток – было выделено 3 промежутка времени, в которых, согласно исследованию Яндекса [13], показатели загруженности дорог отличаются в наибольшую сторону:

- a. День (10:00 – 17:30)
- b. Ночь (19:30 – 08:00)
- c. Часы пиковой загруженности (08:00 – 10:00, 17:30 – 19:30).

3. Дорожное освещение в момент ДТП: присутствует/отсутствует.

Road Defects	
defect_description	COUNT(*)
Не установлены	341k
Отсутствие, плохая различимость горизонтальной разметки проезжей части	121k
Недостатки зимнего содержания	40.8k
Отсутствие дорожных знаков в необходимых местах	40.5k
Отсутствие пешеходных ограждений в необходимых местах	17.2k
Неправильное применение, плохая видимость дорожных знаков	17.1k
Отсутствие освещения	13.8k
Дефекты покрытия	11.0k
Неудовлетворительное состояние обочин	10.1k
Отсутствие тротуаров (пешеходных дорожек)	9.85k
Нарушения в размещении наружной рекламы	5.89k
Отсутствие дорожных ограждений в необходимых местах	4.10k
Низкие сцепные качества покрытия	2.92k

Рисунок 3. Дефекты дорожной среды, повлиявшие на ДТП в России в 2015-2018 гг.

Traffic Violations	
traffic_violation	COUNT(*)
Нет нарушений	793k
Другие нарушения ПДД водителем	91.3k
Несоблюдение очередности проезда	90.7k
Нарушение правил проезда пешеходного перехода	52.9k
Неправильный выбор дистанции	48.5k
Несоответствие скорости конкретным условиям движения	44.6k
Выезд на полосу встречного движения	43.1k
Нарушение правил расположения ТС на проезжей части	39.1k
Переход через проезжую часть вне пешеходного перехода в зоне его видимости либо при наличии в непосредственной близости подземного (надземного) пешеходного перехода	27.0k
Переход через проезжую часть в неустановленном месте (при наличии в зоне видимости перекрёстка)	15.4k
Несоблюдение условий, разрешающих движение транспорта задним ходом	14.5k
Иные нарушения	14.4k
Превышение установленной скорости движения	14.3k

Рисунок 4. Нарушения ПДД, повлиявшие на причины или последствия ДТП

4. Состояние дорожного покрытия в момент ДТП: сухое/мокрое.
5. Тип дорожного покрытия.
6. Наличие перекрестка.
7. Наличие пешеходного перехода.
8. Нарушение правил дорожного движения одним из участников ДТП: Да/Нет.
9. Наличие дорожных дефектов в месте ДТП: Да/Нет.
10. Наличие дефекта в автомобиле, повлиявшем на ДТП: Да/Нет.
11. Тип транспортного средства:
 - a. Легковые автомобили
 - b. Грузовые автомобили, а также автобусы, троллейбусы и т.п.
 - c. Микроавтобусы, минивэны и т.п.

- d. Двухколесные транспортные средства (мотоциклы, велосипеды и т.п.).

К сожалению, в статистике о ДТП в России, предоставляемой ГИБДД, не хватает многих важных показателей, таких как:

1. Разрешенная скорость на участках, в которых произошли ДТП.
2. Наличие светофоров и других дорожных знаков в местах ДТП.
3. Средняя загруженность участка дороги, на котором произошло ДТП.
4. Ширина дороги (число полос).
5. Плотность населения в местах ДТП.
6. Кривая поворота дороги в месте ДТП.

Указанные выше данные можно частично получить из других открытых источников, таких как OpenStreetMap [14] (есть данные об ограничениях скорости, количестве полос на дорогах и т. п.), портала открытых данных РФ [15] и других.

Вывод. Рассмотренные примеры применения алгоритмов искусственного интеллекта в развитых странах показывают перспективность использования подходов анализа данных для улучшения безопасности дорожного движения в России. В рамках работы было разработано программное обеспечение для сбора и визуализации данных о ДТП из открытой статистики ГИБДД. Представлен пример входных данных, которые могут быть использованы для прогнозирования ДТП. Выявлены недостатки в открытых данных, доступных на данный момент. В будущих исследованиях мы планируем собрать большее количество данных из других открытых источников и реализовать на их основе модель прогнозирования ДТП на определенных участках дорог в зависимости от условий окружающей среды, а также анализа исторических данных о ДТП.

ЛИТЕРАТУРА

1. Показатели состояния безопасности дорожного движения (Электронный ресурс - <http://stat.gibdd.ru/>).
2. Global status report on road safety 2018 (Электронный ресурс - https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/).
3. Статистика ДТП в России и мире. Досье (Электронный ресурс - <https://tass.ru/info/3233185>).
4. ГИБДД ищет аварийный выход (Электронный ресурс - <https://www.kommersant.ru/doc/3398213>).
5. Tennessee Highway Patrol (Электронный ресурс - [https://www-01.ibm.com/common/ssi/cgi-](https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?subtype=AB&infotype=PM&htmlfid=GVC03014USEN&attachment=GVC03014USEN.PDF)
- bin/ssialias?subtype=AB&infotype=PM&htmlfid=GVC03014USEN&attachment=GVC03014USEN.PDF).
6. Evaluating road safety and safety effects using Empirical Bayesian method (Электронный ресурс - <https://www.itf-oecd.org/sites/default/files/docs/8-peltola.pdf>).
7. A review of accident prediction models for road intersections (Электронный ресурс - https://www.researchgate.net/publication/265108102_A_review_of_accident_prediction_models_for_road_intersections).
8. Анализ причин и следствий дорожно-транспортных происшествий (Электронный ресурс - http://statsoft.ru/solutions/ExamplesBase/tasks/detail.php?ELEMENT_ID=702).
9. Повышение эффективности прогнозирования ДТП на автомобильных дорогах вне населенных пунктов на основе разработки экспертной системы (Электронный ресурс - <http://izvestia.asu.ru/ru/article/702/>).
10. Яндекс разработал систему прогнозирования пробок и ДТП (Электронный ресурс - https://yandex.ru/company/services_news/2015/0302).
11. Наука о пробках (Электронный ресурс - <https://rg.ru/2017/07/04/reg-cfo/aleksandr-poliakov-my-nauchilis-prognozirovat-dtp.html>).
12. Apache Superset (incubating) (Электронный ресурс - <https://superset.incubator.apache.org/>).
13. Автомобильные пробки в Санкт-Петербурге (Электронный ресурс - <https://yandex.ru/company/researches/2015/spb/jams>).
14. OpenStreetMap (Электронный ресурс - <https://www.openstreetmap.org>).
15. Портал открытых данных Российской Федерации (Электронный ресурс - <https://data.gov.ru/>).

REVIEW OF APPROACHES TO THE APPLICATION OF ALGORITHMS OF ARTIFICIAL INTELLIGENCE FOR DETECTING FACTORS AFFECTING THE RISK OF ROAD ACCIDENTS OCCURENCE

© 2018 D. S. Donchenko, N. P. Sadovnikova, D. S. Parygin

Volgograd State Technical University (Volgograd, Russia)

The paper describes various approaches to the application of artificial intelligence algorithms for analyzing data on road traffic accidents in different countries of the world. Proposed an approach to use of artificial intelligence to analyze data on road accidents in Russia in order to detect the factors that increase the risk of road accidents.

Key words: machine learning, artificial intelligence, data analysis.

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-47-340012 р_а, 18-37-20066 мол_а_вед.