

ISBN 978-83-959742-6-7

Transport Problems 2023

Conference proceedings

Konferencja dofinansowana ze środków budżetu państwa w ramach programu Ministra Edukacji i Nauki pod nazwą "Doskonała Nauka" nr projektu DNK/SP/549716/2022
kwota dofinansowania 145 200,00 zł
całkowita wartość projektu 299 000,00 zł.



Ministerstwo
Edukacji i Nauki



Doskonała
Nauka

UNDER THE HONORARY
PATRONAGE OF



Silesian
University
of Technology



Krajowa
Reprezentacja
Doktorantów

KRD



XV INTERNATIONAL
SCIENTIFIC
CONFERENCE

21-23.06 2023
Katowice - Częstochowa

19-20.06 2023
Katowice

XII INTERNATIONAL
SYMPOSIUM OF YOUNG
RESEARCHERS



Silesian University of Technology
Faculty of Transport and Aviation Engineering

Transport Problems 2023

Proceedings

XV International Scientific Conference

XII International Symposium of Young Researchers

UNDER THE HONORARY PATRONAGE OF



Silesian University
of Technology



Krajowa
Reprezentacja
Doktorantów

KRD

Media patronage:

Transport Problems International Scientific Journal
ISSN 1896-0596, The Silesian University of Technology,
Faculty of Transport and Aviation Engineering

Transport Problems

International Scientific Journal



Ministerstwo
Edukacji i Nauki



Doskonała
Nauka

Konferencja dofinansowana ze środków budżetu państwa
w ramach programu Ministra Edukacji i Nauki pod nazwą
"Doskonała Nauka" nr projektu DNK/SP/549716/2022
kwota dofinansowania 145 200,00 zł
całkowita wartość projektu 299 000, 00 zł.

editor-in-chief

A. Sładkowski

editorial board

*P. Czech, M. Cieśla, T. Haniszewski,
M. Juzek, W. Kamiński, P. Marzec, G. Wojnar*



XII INTERNATIONAL SYMPOSIUM OF YOUNG RESEARCHERS TRANSPORT PROBLEMS 2023

TABLE OF CONTENTS

No.	Author, title	Pages	
		Begin	End
1	Mujahid ALI, Elżbieta MACIOSZEK Travel mode choice, multi-tasking activities, and its influence on subjective well-being	652	662
2	Kamil BOBIŃSKI, Adam RAJCA, Michał BANET, Robert WIESZAŁA, Henryk BĄKOWSKI Cessna winglet optimization - a comparative study of different designs	663	667
3	Oussama BOUGUEFFA, Patryk TEZCAN, Kundai BGONI, Aleksander ŚLADKOWSKI The use of automated vehicles in warehouses	668	677
4	Michał BRZOZOWSKI, Kacper CIEŚLAR, Krzysztof PARCZEWSKI The effect of snow on the operation of autonomous car sensors	678	686
5	Artur BUDZYŃSKI, Aleksander ŚLADKOWSKI, Bartłomiej DŁUGOSZ, Konrad SIKORA, Wesam HASAN, Taras NIKITISHYN The analysis of factors affecting the occurrence of road incidents in Poland in 2015-2021	687	695
6	Martynas DRAZDAUSKAS Numerical studies of ice combustion cycle operating on dual ammonia diesel fuel	696	709
7	Krzysztof DREWNIOK, Aleksander SOBOTA Test results of inter-stops travel times on bus line no. 11 from Katowice to Czeladz	710	719
8	Mikołaj FEDEROWICZ, Dawid GOŁAŚ, Adam JABŁOŃSKI, Izabella KOSTORZ, Kamil ULMAN, Krystian SZEWCZYŃSKI, Piotr OPAŁA Project of using a modern car simulator to analyze road safety in selected urban areas	720	725
9	Isidore KACHAKHIDZE, Omar KIKVIDZE Numerical calculation of vibrations of a single-axis trailer	726	730

Keywords: traffic accidents; big data; forecasting

Artur BUDZYŃSKI*, Aleksander ŚLADKOWSKI, Bartłomiej DŁUGOSZ, Konrad SIKORA, Wesam HASAN

Silesian University of Technology
Kraśińskiego 8, 40-019 Katowice, Poland

Taras NIKITISHYN

Academic Secondary School of the Silesian University of Technology
Księża Marcina Strzody 10, 44-100 Gliwice, Poland

*Corresponding author. E-mail: artur.budzynski@polsl.pl

THE ANALYSIS OF FACTORS AFFECTING THE OCCURRENCE OF ROAD INCIDENTS IN POLAND IN 2015-2021

Summary. Accidents and collisions are undesirable. The article presents an analysis of dependencies affecting the occurrence of an accident or collision. Data for the study come from the Polish Police system. Over 3 million accidents and collisions have been analysed. A machine learning model was trained to classify accidents and collisions. Modern tools were used: Python programming language, Jupyter Notebook for experiments. Programming libraries were used: Pandas for data processing, NumPy for calculations, Matplotlib for visualization and Scikit-Learn for training the machine learning model.

ANALIZA CZYNNIKÓW WPŁYWAJĄCYCH NA WYSTĘPOWANIE ZDARZEŃ DROGOWYCH W POLSCE W LATACH 2015-2021

Streszczenie. Wypadki i kolizje są niepożądane. W artykule przedstawiono analizę zależności wpływających na wystąpienie wypadku lub kolizji. Dane do badania pochodzą z systemu polskiej Policji. Przeanalizowano ponad 3 miliony wypadków i kolizji. Wyszkolono model uczenia maszynowego w celu klasyfikowania wypadków i kolizji. Wykorzystano nowoczesne narzędzia: język programowania Python, Jupyter Notebook do eksperymentów. Wykorzystano biblioteki programistyczne: Pandas do przetwarzania danych, NumPy do obliczeń, Matplotlib do wizualizacji oraz ScikitLearn do trenowania modelu uczenia maszynowego.

1. WSTĘP

Wypadki drogowe są niepożądane. Problem dotyczy całego świata. Wypadki są przyczyną śmierci i obrażeń wielu ludzi. Wypadki drogowe są wynikiem splotu czynników. Tych czynników jest wiele. Przyczyny wypadków można podzielić na bezpośrednie i pośrednie. Istnieją czynniki ludzkie, techniczne i zewnętrzne. Czynniki ludzkie popełnia błąd. Czynniki pośrednie wpływają na ludzkie błędy. Zmęczenie jest pośrednią przyczyną błędów ludzkich. Brak zmęczenia może być czynnikiem, który może zapobiec wypadkowi spowodowanemu przez czynnik zewnętrzny lub techniczny.

W [1] przeprowadzono badania na ludziach z Południowej Walii. Badanie przeprowadzono za pomocą kwestionariuszy. Sprawdzono zależność udziału w wypadku od spożycia kofeiny. Wskaźnik wypadków wśród osób spożywających kofeinę wyniósł 1,9%, a w grupie niespożywającej kofeiny

3,3%. Bardziej narażone na udział w wypadkach drogowych są osoby młodsze, w złym stanie zdrowia, zestresowane i podejmujące większe ryzyko.

Wypadki w Sudanie analizowano w [2]. Dane pochodziły z kwestionariuszy sporządzonych przez ofiarę wypadku i naocznego świadka. Wykorzystano również raporty policyjne. Wypadek miał wpływ geograficzny. Najwięcej wydarzyło się w stolicy kraju - Chartumie 46,7%. Czynnikiem wpływającym było miejsce zdarzenia. Wypadki na przejazdach stanowiły 46,8% ogółu. Przyczyną wypadków były błędy kierowców. Jednym z błędów jest nadmierna prędkość. Jest to świadomy błąd, który odróżnia go od pozostałych, odpowiadając za 14,6% wypadków.

Rys. 1 przedstawia cykl życia modelu uczenia maszynowego. Praca nad modelem to proces ciągłego doskonalenia. Zaczyna się od wyboru źródeł danych. Jakość danych może być często ważniejsza niż sama metoda używana do uczenia modelu. Kolejne etapy to przygotowanie i przekształcenie danych. Duże zbiory danych są często przedstawione w sposób nieidealny. Wymagają przygotowania do formy potrzebnej do analizy. Wymaga to poza umiejętnościami programistycznymi wiedzy branżowej. Kolejne etapy to analiza i praca nad właściwym modelem do prognozowania. Ostatnim etapem jest wdrożenie. Jest to etap, którego nie przechodzi większość modeli.

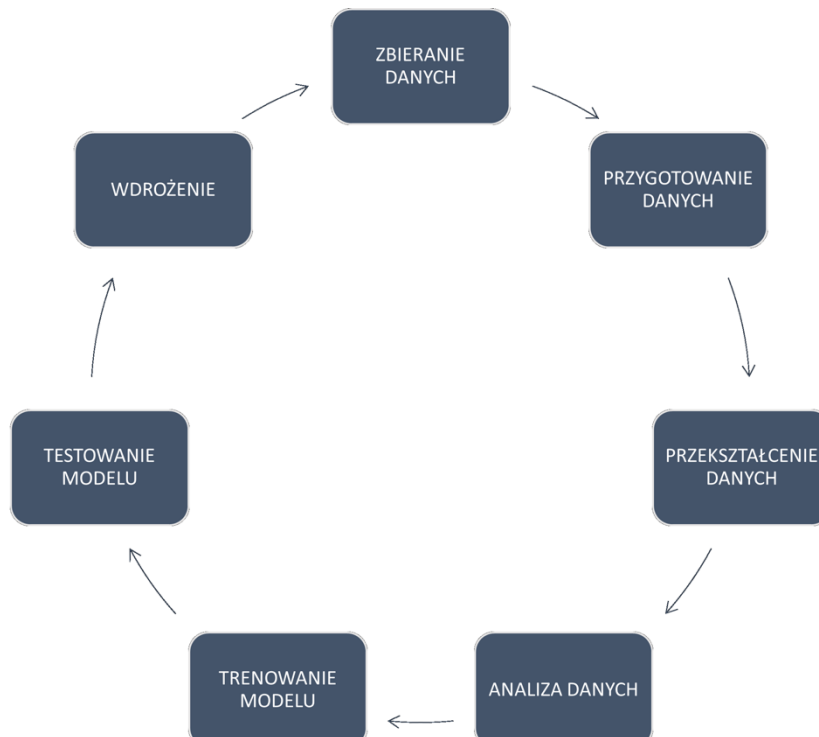


Fig. 1. Machine learning model life cycle

Rys. 1. Cykl życia modelu uczenia maszynowego

Rys. 2 przedstawia liczbę cytowań artykułów poruszających tematykę oprogramowania do pracy z uczeniem maszynowym i dużymi zbiorami danych. Powszechnie stosowaną praktyką jest udostępnianie wartościowych rozwiązań w otwartym dostępie. To oprogramowanie jest również bezpłatne. Autorzy publikują artykuły, które osoby korzystające z nich w swoich pracach badawczych cytują. Omówione w artykułach narzędzia mają zastosowanie w rozwiązaniu wielu problemów nauki, przemysłu i biznesu.

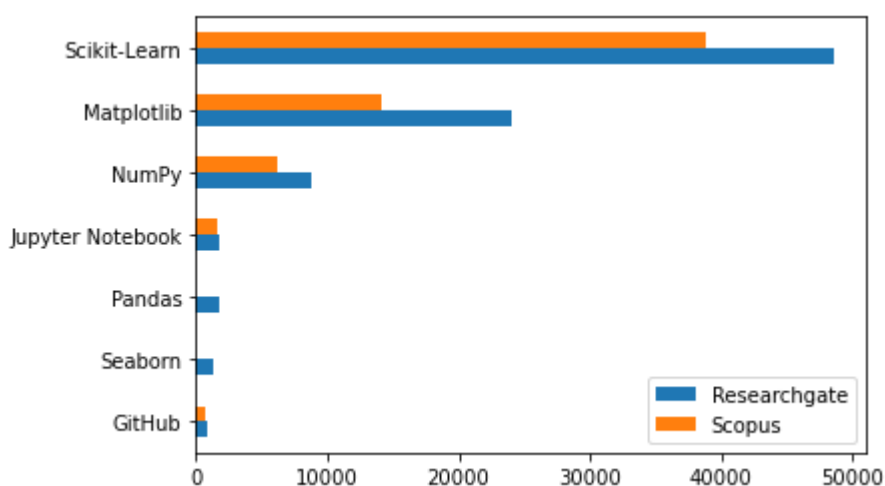


Fig. 2. Scopus and Researchgate citations
Rys. 2. Cytowania w bazie Scopus i Researchgate

2. METODYKA

Projekt został napisany przy użyciu języka programowania Python [3]. Kod jest napisany za pomocą Jupyter Notebook [4]. Jest to interaktywne środowisko internetowe do pisania i udostępniania kodu. Do przetwarzania danych wykorzystano bibliotekę Pandas [5]. Tworzenie nowych cech odbywało się poprzez sumowanie wartości kolumn w Data Frame Pandas. Do wizualizacji danych wykorzystano biblioteki Matplotlib [6] oraz Seaborn [7]. Do tworzenia wykresów słupkowych i liniowych użyto Matplotlib. Seaborn został wykorzystany do stworzenia mapy cieplnej.

Dane wykorzystane do opracowania dotyczą wypadków, które miały miejsce na terenie Polski w latach 2017-2021. Dane zostały otrzymane od Komendy Głównej Policji. Dane dotyczą ponad 3 milionów wypadków i kolizji drogowych.

Powołując się na definicję terminu „poszkodowany” Głównego Urzędu Statystycznego. Osoba poszkodowana to osoba, która nie zginęła na miejscu w wyniku wypadku lub zmarła w ciągu 30 dni, ale odniosła obrażenia, które zazwyczaj wymagają leczenia. Osoba ciężko ranna to osoba poszkodowana, która w wyniku nieszczęśliwego wypadku przebywała w szpitalu dłużej niż 24 godziny. Lekko ranny to osoba ranna, z wyłączeniem zabitych i ciężko rannych. Osoby, które są mniej ranne, takie jak drobne skaleczenia lub siniaki, zwykle nie są rejestrowane jako ranne. Na podstawie tych informacji istnieje cienka granica między sklasyfikowaniem osoby jako ciężko rannej lub martwej. Na przykład osoba, która umiera 31 dnia po wypadku, jest statystycznie ciężko ranna. Istnieje cienka granica między zaklasyfikowaniem osoby jako ciężko rannej lub lekko rannej. Po wypadku uczestnikom wypadku towarzyszy wiele emocji związanych ze stresem. Osoba, która zaraz po zdarzeniu czuje się dobrze i nie wymaga pomocy medycznej, może czuć się gorzej, gdy emocje przestaną działać. Na potrzeby projektu postanowiono obliczyć zmienną dodatkową „ofiary” jako sumę zabitych, ciężko rannych i lekko rannych.

3. REZULTATY

Przeprowadzono analizę z wykorzystaniem narzędzi opisanych w przeglądzie literatury. Rys. 3 przedstawia podstawowe informacje o zbiorze danych. Wykonano to z wykorzystaniem funkcji `df.info()`. IDKSIP zdarzenia i ID zdarzenia to oznaczenia zdarzenia. Data zdarzenia przedstawia datę i godzinę zdarzenia. Godzina jest podana z dokładnością do minuty. Powiat, Gmina, Województwo, Miejsowość, Ulica, Numer domu, opisują adres i oznaczają dokładnie co to nazwa kolumny. Numer

drogi opisuje numer drogi lub podaje informację jakiej jest klasy. KM HM (Pikietaż) podaje informację o umiejscowieniu, jest to informacja widoczna na czerwono białych słupkach przy drogach. Rondo/węzeł, Skrzyżowanie z ulicą, Skrzyżowanie z drogą podaje nazwę, jeśli jej nie ma wpisane jej NaN. Odległość do skrzyżowania podaje odległość w metrach. GPS x i GPS y opisują koordynaty GPS. Wypadek/Kolizja i Obszar to binarne wybory.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3115047 entries, 0 to 3115046
Data columns (total 32 columns):
#   Column                                     Dtype
---  -
0   IDKSIP zdarzenia                          object
1   ID zdarzenia                               int64
2   Data zdarzenia                            object
3   Powiat                                     object
4   Gmina                                      object
5   Wojewodztwo                               object
6   Miejscowość                               object
7   Ulica                                       object
8   Numer domu                                object
9   Numer drogi                               object
10  KM HM (Pikietaż)                           object
11  Rondo/węzeł                               object
12  Skrzyżowanie z ulicą                       object
13  Skrzyżowanie z drogą                       object
14  Odległość do skrzyżowania                  float64
15  GPS x                                       object
16  GPS y                                       object
17  Wypadek/Kolizja                           object
18  Obszar                                     object
19  Geometria drogi                            object
20  Skrzyżowanie                               object
21  Charakterystyka miejsca zdarzenia          object
22  Oświetlenie                                object
23  Warunki atmosferyczne                      object
24  Rodzaj zdarzenia                           object
25  Prędkość dopuszczalna                      float64
26  Rodzaj drogi                               object
27  Nawierzchnia drogi                         object
28  Stan nawierzchni                           object
29  Sygnalizacja świetlna                     object
30  Oznakowanie poziome                        object
31  Inne przyczyny                             object
dtypes: float64(2), int64(1), object(29)
memory usage: 760.5+ MB
```

Fig. 3. Key information about the dataset

Rys. 3. Kluczowe informacje o zbiorze danych

Badanie opisane w [8] miało na celu zbadanie względnego ryzyka wypadku w przypadku różnych drogowych warunków pogodowych i ich kombinacji. W badaniu zastosowano niedawno opracowaną metodę opartą na pojęciu prawdopodobieństwa Palma, wywodzącym się z teorii losowych procesów punktowych. Badaniem objęto wszystkie zgłoszone przez policję wypadki jedno - i wielopojazdowe mające miejsce na 43 głównych drogach w Finlandii w latach 2014–2016. Głównym wkładem pracy jest zademonstrowanie metody na skalę krajową poprzez wykorzystanie szacunkowych godzinowych natężeń ruchu na odcinkach dróg zamiast zmierzonych, które byłyby dostępne tylko dla nielicznych dróg. Porównując względne ryzyko wypadku w zależności od typu drogi, wyniki wykazały, że ryzyko przy złych warunkach pogodowych i drogowych było wyższe na autostradach w porównaniu z drogami dwupasmowymi i wielopasmowymi, chociaż ogólne ryzyko było niższe na autostradach. ryzyko wypadku było generalnie wyższe w przypadku wypadków z udziałem jednego pojazdu w porównaniu z wypadkami z udziałem wielu pojazdów. Liczba wypadków zależy od pogody. Rys. 4 przedstawia liczbę wypadków latach 2015-2020 w Polsce w zależności od pogody. Najwięcej wypadków było przy dobrej pogodzie a zdecydowanie mniej przy utrudnieniach pogodowych. Może to być spowodowane zbyt dużą złągubną pewnością siebie kierowców w dobrych warunkach. Utrudnienia mogą powodować większe skupienie.

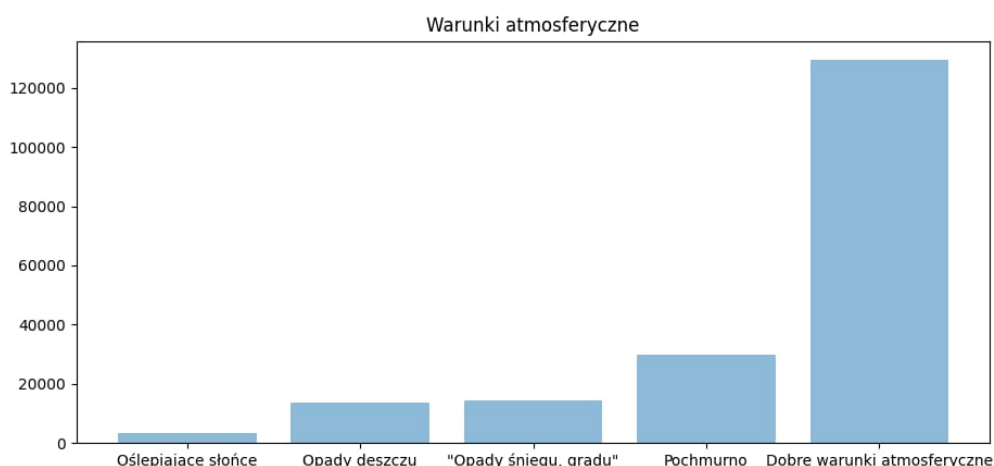


Fig. 4. Number of accidents depending on the weather
Rys. 4. Ilość wypadków w zależności od pogody

Liczba wypadków zależy od stanu nawierzchni. Rys. 5 przedstawia liczbę wypadków w Polsce w latach 2015-2020 według stanu nawierzchni. Ten wykres jest bardzo zbliżony do poprzedniego. Naturalną konsekwencją opadów jest mokra nawierzchnia.

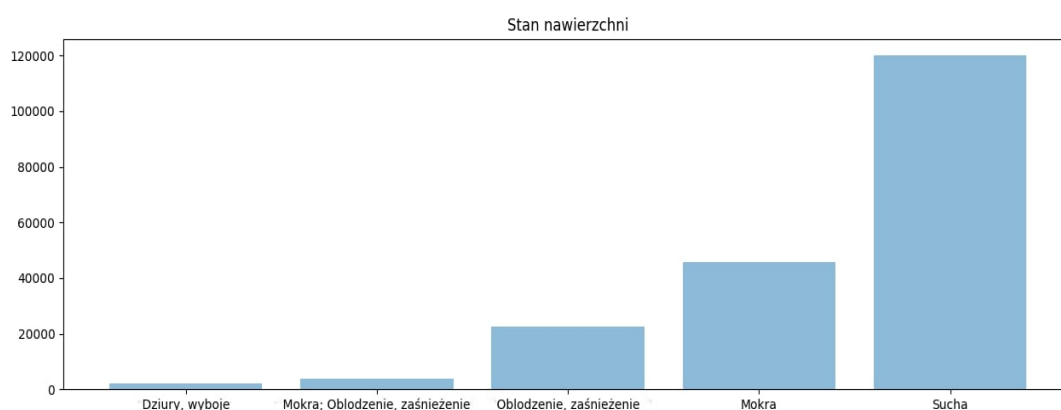


Fig. 5. Number of accidents depending on the condition of the road surface
Rys. 5. Ilość wypadków w zależności od stanu nawierzchni

Ilość wypadków zależy od dnia tygodnia. Rys.6 pokazuje liczbę wypadków w latach 2015-2020 w Polsce w podziale na dni tygodnia. Wykres można wytłumaczyć zwiększoną aktywnością w dni robocze ze względu na dojazdy do miejsca edukacji lub pracy. Aktywność ta w dni wolne od pracy jest mniejsza.

Liczba zdarzeń zależy od czasu ich wystąpienia. Rys. 7 przedstawia liczbę zdarzeń z podziałem na godziny. Ciemna niebieska linia pokazuje średnią wypadków o danej godzinie. Jasny ślad pokazuje zakres między minimum a maksimum. Ilość zdarzeń rośnie w godzinach od 3:00 do 16:00 osiągając największą wartość. Od 16:00 zaczyna spadać. Zależność jest związana z popołudniowym szczytem komunikacyjnym.

Liczba przypadków zależy od oświetlenia. Rys. 8 przedstawia przyczyny wypadków w Polsce w latach 2015-2020 według oświetlenia. Wykres potwierdza analizę z poprzedniego. Łatwo można powiązać godziny wystąpienia z naturalnym oświetleniem.

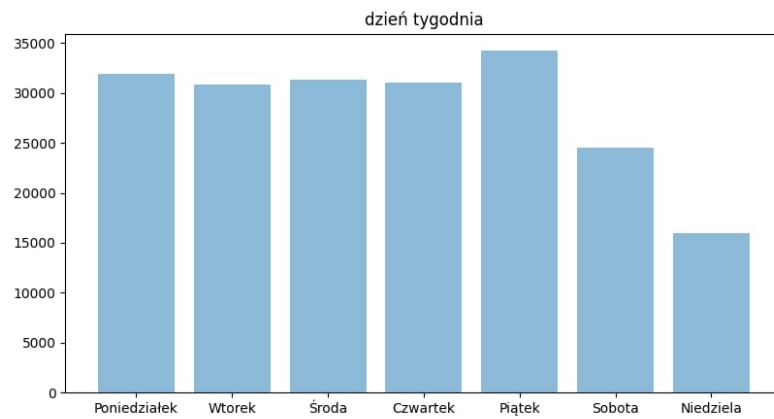


Fig. 6. Number of accidents by day of the week
Rys. 6. Ilość wypadków z podziałem na dni tygodnia

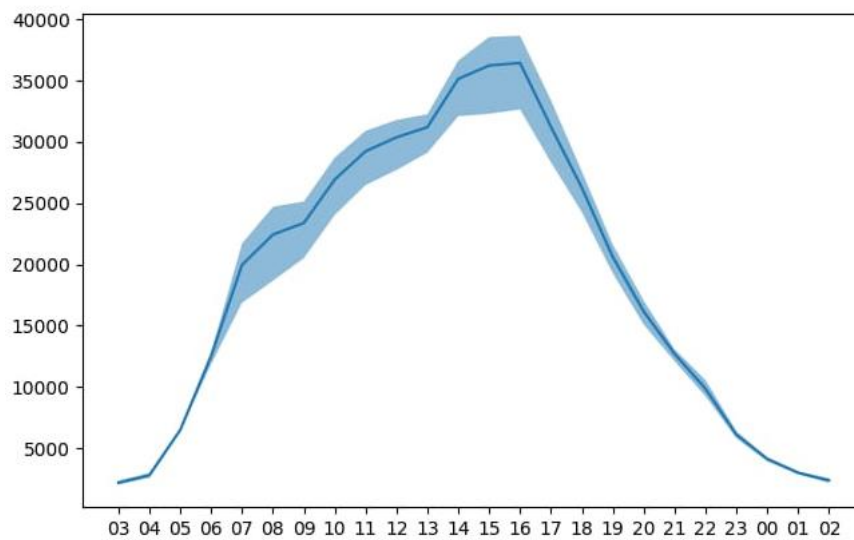


Fig. 7. Dependence of the occurrence of the accident on the hour
Rys. 7. Zależność wystąpienia wypadku od godziny

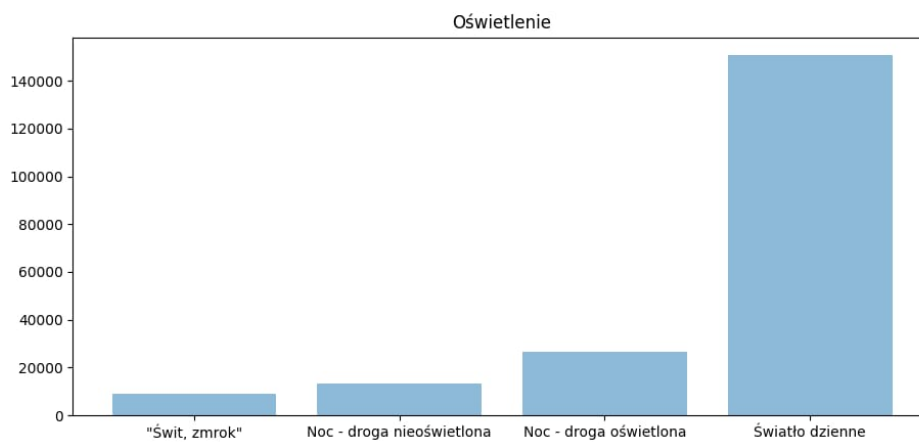


Fig. 8. Number of accidents depending on the lighting
Rys. 8. Ilość wypadków w zależności od oświetlenia

Przyczyny zdarzeń drogowych są zróżnicowane. Rys. 9 przedstawia wystąpienia zdarzeń drogowych z podziałem ze względu na przyczynę. Najwięcej zdarzeń jest bocznych. Jest to spowodowane słabą ochroną człowieka od boku pojazdu w porównaniu z przodem. Przód pojazdu ma strefę zgniotu, która zmniejsza negatywny wpływ zderzenia na człowieka. Sporą część stanowią najechania na pojazd unieruchomiony. Mogą to być zarówno pojazdy stojące w korku lub na poboczu. Warto tu zwrócić uwagę na odpowiednie oznaczenie pojazdów znajdujących się na poboczu podczas awarii.

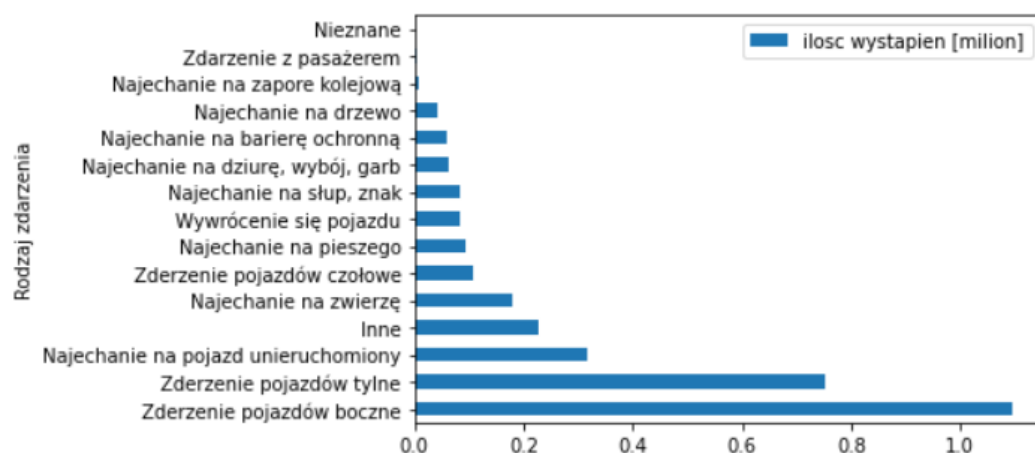


Fig. 9. Number of events depending on the type

Rys. 9. Ilość zdarzeń w zależności z podziałem na rodzaj

Rys. 10 przedstawia wykres wystąpień zdarzeń drogowych z podziałem na rodzaj drogi. Zdecydowanie największy udział mają drogi dwukierunkowe jedno jezdniowe. Na nich dochodzi do bardzo wielu niebezpiecznych zdarzeń. Sytuacja w Polsce zmienia się na dobre w ostatnich latach. Powstało wiele dróg dwu jezdniowych dwukierunkowych, ekspresowych i autostrad. Kraj można przejechać w wielu miejscach od granicy do granicy z ominięciem kolizyjnych skrzyżowań. Transport towarowy towarów został w wielu miejsca wyprowadzony z miast. Powoduje to poprawę bezpieczeństwa i pozwala stwierdzić, że możliwe jest zmniejszenie ilości zdarzeń drogowych z latami kolejnych.

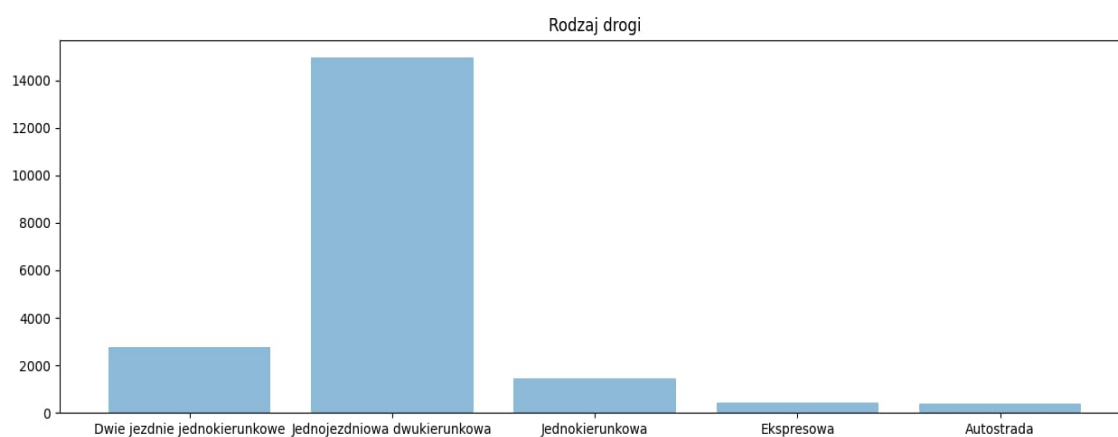


Fig. 10. Number of events broken down by road type

Rys. 10. Ilość zdarzeń z podziałem na rodzaj drogi

Wypadki są zależne od miejsca. Przez miejsce można rozumieć typ infrastruktury jak w poprzednim wykresie oraz geograficznym. Rys. 11 przedstawia wykres zależności ilości wystąpienia wypadków od województwa. Najwięcej wypadków wystąpiło w województwie mazowieckim. Jest to duże województwo ze stolicą kraju. W województwie mazowieckim występuje też ruch tranzytowy.

Niemalże 400 000 wypadków wystąpiło w województwie śląskim. Jest to województwo bardzo rozwinięte.

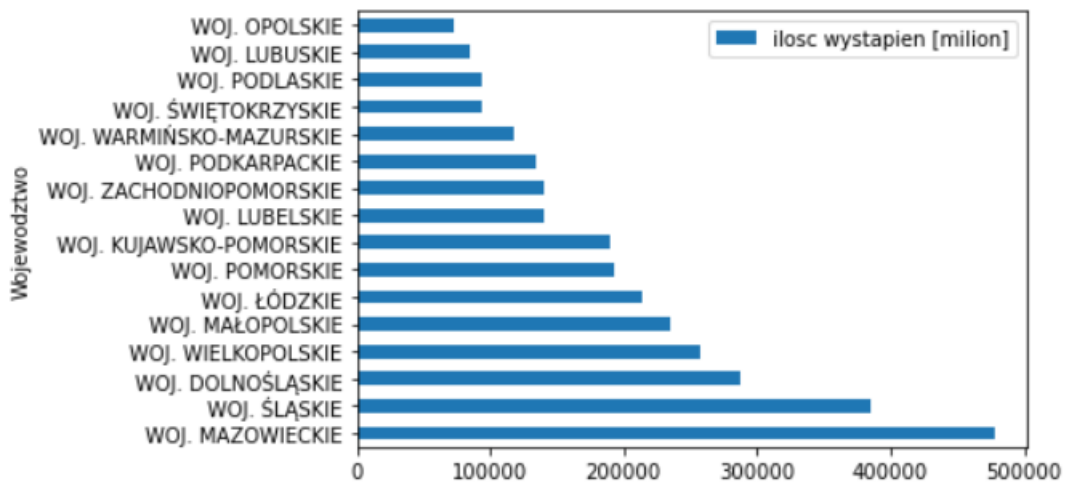


Fig. 11. Number of events broken down by road type

Rys. 11. Ilość zdarzeń z podziałem na rodzaj drogi

Zaproponowano wytrenowanie modelu uczenia maszynowego do klasyfikacji między wypadkiem a kolizją. Porównano kilka modeli mających zastosowanie w podobnych problemach. Wybrano GradientBoostingClassifier z największą dokładnością 93,8. Wyniki przedstawiono na rys. 12. Najważniejszą cechą było rodzaj zdarzenia. Inne cechy są mniej istotne dla modelu.

Weight	Feature
0.7219 ± 0.5255	Rodzaj zdarzenia
0.0778 ± 0.3091	Charakterystyka miejsca zdarzenia
0.0330 ± 0.2866	Województwo
0.0214 ± 0.1254	Numer drogi
0.0212 ± 0.2765	Powiat
0.0198 ± 0.1323	Inne przyczyny
0.0191 ± 0.0406	Ulica
0.0132 ± 0.0463	Skrzyżowanie
0.0112 ± 0.0839	Gmina
0.0086 ± 0.0648	Rok
0.0085 ± 0.1137	Dzień roku
0.0066 ± 0.0494	Obszar
0.0065 ± 0.0550	Prędkość dopuszczalna
0.0063 ± 0.0601	Rodzaj drogi
0.0052 ± 0.0345	Stan nawierzchni
0.0045 ± 0.0569	Oznakowanie poziome
0.0035 ± 0.0364	Oświetlenie
0.0032 ± 0.0458	Dzień tygodnia
0.0023 ± 0.0649	Numer domu
0.0017 ± 0.0754	Warunki atmosferyczne
	... 9 more ...

Fig. 12. The most important features of the model

Rys. 12. Najważniejsze cechy modelu

4. WNIOSKI

Wykorzystanie powszechnej rejestracji danych o wypadkach przez siły policyjne na całym świecie pozwoli na lepsze porównywanie danych między krajami. Łatwiejszym do wdrożenia rozwiązaniem jest zastosowanie uniwersalnej metody ewidencjonowania danych o wypadkach dla policji w Unii

Europejskiej. Zaleca się, aby siły policyjne na całym świecie miały otwarty dostęp do danych w celu skonsolidowania wysiłków badawczych. Liczba wypadków spada z roku na rok.

Wiele przydatnych narzędzi do pracy z dużymi zbiorami danych i uczeniem maszynowym jest dostępne w otwartym dostępie. Prace naukowe dotyczące tych programów są bardzo często cytowane. Omówione w przeglądzie literatury oprogramowanie znalazło zastosowanie w analizie zbioru danych dotyczącego wypadków.

Modelowanie wypadku pozwala lepiej zrozumieć zależności, które wpływają na jego wystąpienie. Wypadki są absolutnie negatywne, a wysiłki zmierzające do ich ograniczenia i zmniejszenia wszelkich negatywnych skutków są cenne. Zdarzenia drogowe występują częściej przy dobrych warunkach pogodowych i suchej nawierzchni. Najwięcej wypadków jest w piątki. Może to być spowodowane zwiększoną aktywnością na drogach. Taką aktywność można powiązać ze zmianą miejsca na weekend. Mniej wypadków jest w dni wolne niż robocze. Ilość wypadków rośnie od ok godziny 3 do ok godziny 16. Od ok godziny 16 do 3 spada. Najniebezpieczniejsze są zderzenia pojazdów boczne. Zderzenia boczne są szczególnie niebezpieczne ze względu na stosunkowo cienkie w porównaniu do przodu pojazdu drzwi. Do największej ilości wypadków dochodzi na drogach jedno jezdniowych dwukierunkowych.

Dane dla tego projektu są dostępne w otwartym dostępie.

4. PODZIĘKOWANIA

Projekt realizowany w ramach VIII konkursu finansowania kształcenia zorientowanego projektowo – PBL, w ramach programu Inicjatywa Doskonałości – Uczelnia Badawcza Politechniki Śląskiej.

Bibliografia

1. Smith, A. Caffeine and road traffic accidents. *World journal of pharmaceutical and medical research*. 2021. Vol. 7. P. 1-4.
2. Abeer Hassan, A. & Osman, A. & Alhassan, A. & Hussain, G. & Hafeez Ahmed, A. & Osman, M. & Ahmed, H. & Elhassan, A. Epidemiology of road traffic accidents in Sudan. *European Academic Research*. 2016. Vol. 3. No. 11. P. 12285-12299.
3. Van Rossum, G. & Drake, F.L. *Python 3 Reference Manual*. 2009.
4. Kluyver, T. & Ragan-Kelley, B. & Perez, F. & Granger, B. & Bussonnier, M. & Frederic, J. & Kelley, K. & Hamrick, J. & Grout, J & Corlay, S. & Ivanov, P. & Avila, D & Abdalla, S. *Jupyter notebooks – a publishing format for reproducible computational workflows*. Positioning and Power in Academic Publishing: Players, Agents and Agendas. 2016.
5. McKinney, W. Data structures for statistical computing in python. *Proc. Of the 9th python in science conference*. 2010. P. 56-61.
6. Hunter, J.D. Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*. 2007. Vol. 9. No.3. P. 90-95.
7. Waskom, M.L. Seaborn: statistical data visualization. *Journal of Open Source Software*. 2021. Vol. 6. No. 60. P. 1-4.
8. Malin, F. & Norros, I. & Innamaa, S. Accident risk of road and weather conditions on different road types. *Accident Analysis & Prevention*. 2019. Vol. 122. P. 181-188.