

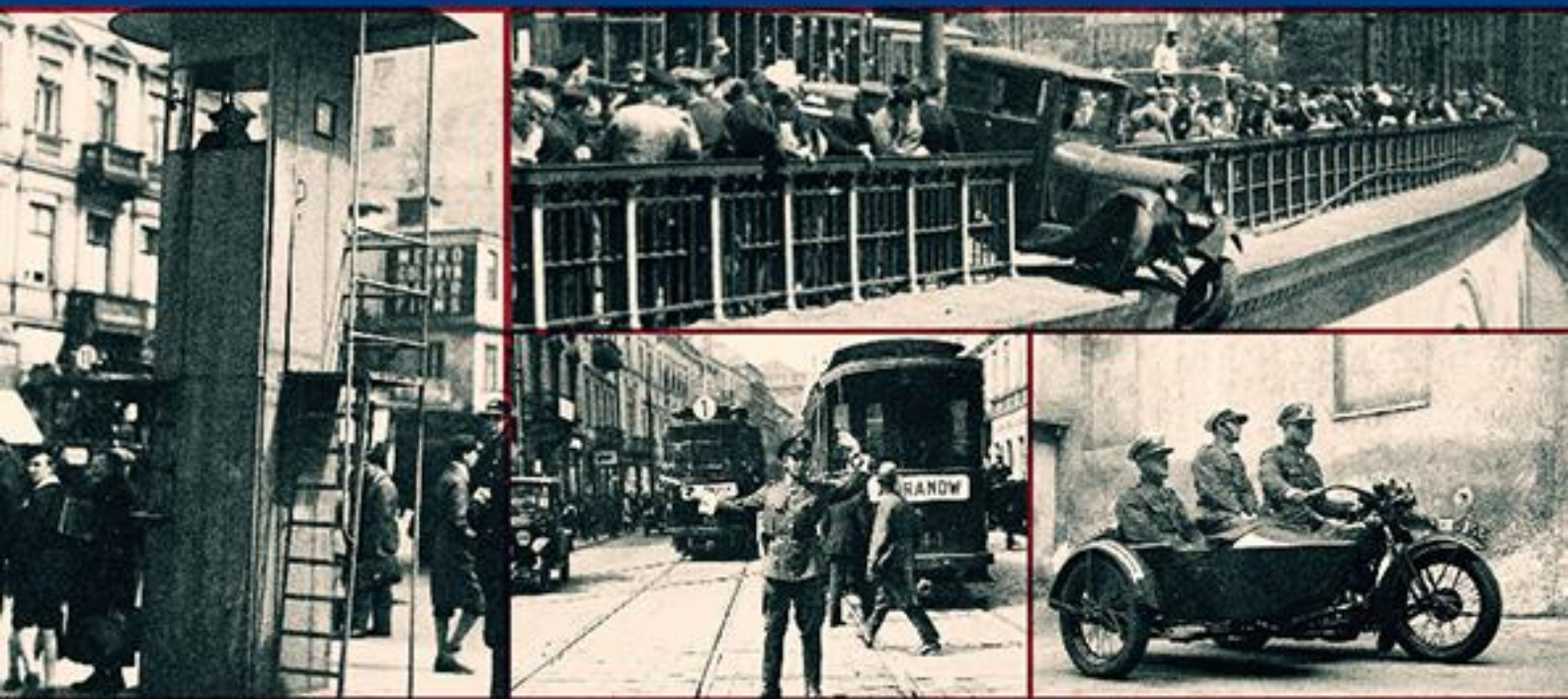
# Kwartalnik policyjny

Nr 1-2(48-49)/2019

Rok XIII

ISSN 1898-1453

CZASOPISMO  
CENTRUM SZKOLENIA POLICJI  
W LEGIONOWIE



## 100 lat NA POSTERUNKU



### W numerze:

Policja Państwowa ■ Policja Województwa Śląskiego ■ Generalowie Policji ■ Kompanie rezerwy PP  
■ Kynologia policyjna ■ Narodziny „drogówki” w Polsce ■ Przewidywanie wypadków komunikacyjnych  
za pomocą sztucznej inteligencji ■ Udział kobiet w misjach zagranicznych



# PRZEWIDYWANIE RYZYKA WYSTĄPIENIA WYPADKÓW KOMUNIKACYJNYCH

## ZA POMOCĄ UCZENIA MASZYNOWEGO (SZTUCZNEJ INTELIGENCJI)

**mł. asp. Krzysztof Kinel**

Specjalista Wydziału Ruchu Drogowego  
Komendy Wojewódzkiej Policji w Katowicach

W ciągu ostatnich 3 lat (2016–2018) na drogach województwa śląskiego zginęło w sumie 725 osób, a ponad 12 tys. zostało rannych<sup>1</sup>. W analizowanym okresie liczba zabitych zmniejszyła się o 14,3%, liczba rannych spadła o 13,6%, z kolei liczba wypadków uległa redukcji o 12,1%. Pomimo iż liczba zdarzeń drogowych, w których uczestnicy odnieśli obrażenia lub ponieśli śmierć, maleje z roku na rok, należy dołożyć wszelkich starań, żeby osiągnąć ostateczny cel: „wizję zero” – zero osób ciężko rannych i ofiar śmiertelnych w wypadkach drogowych.

Z pomocą przychodzą nowe technologie. Uczenie maszynowe jest to dział informatyki, który wyewoluował w połowie XX wieku z badań nad sztuczną inteligencją. Dziedzina ta zajmuje się teorią i praktycznym zastosowaniem algorytmów analizujących dane. Zadaniem uczenia maszynowego jest nauczenie modelu na zbiorze treningowym, aby móc na nowym zbiorze przewidywać odpowiedzi<sup>2</sup>.

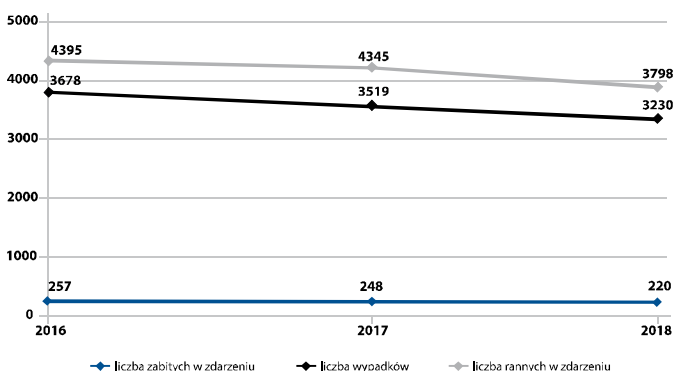
Uczenie maszynowe może pomóc w przewidywaniu ryzyka występowania wypadków drogowych, z wykorzystaniem danych pochodzących z przeszłości i typowania miejsc, w których

prawdopodobieństwo wystąpienia przedmiotowego zdarzenia jest największe. Tym samym zyskuje się dodatkowe narzędzie wspomagające pracę komórek ruchu drogowego.

### ETAPY PRACY Z DANYMI<sup>3</sup>



**Wykres 1.** Stan bezpieczeństwa w ruchu drogowym na terenie województwa śląskiego.



### Pobieranie danych

Na potrzeby przedmiotowego projektu pobrano dane z Systemu Ewidencji Wypadków i Kolizji z lat 2016, 2017 i 2018<sup>4</sup>. Struktura danych, które wykorzystano, wygląda następująco: współrzędne geograficzne, miejscowość, ulica, numer drogi, liczba wypadków drogowych, liczba osób rannych, liczba osób zabitych.

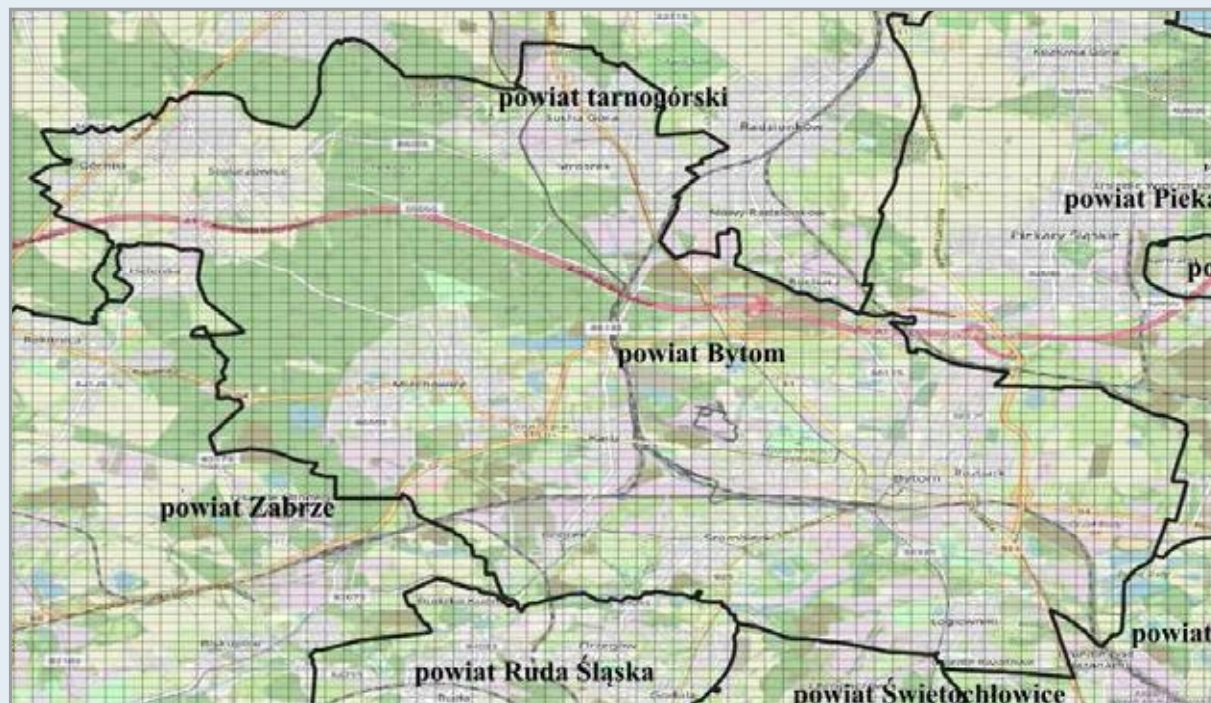
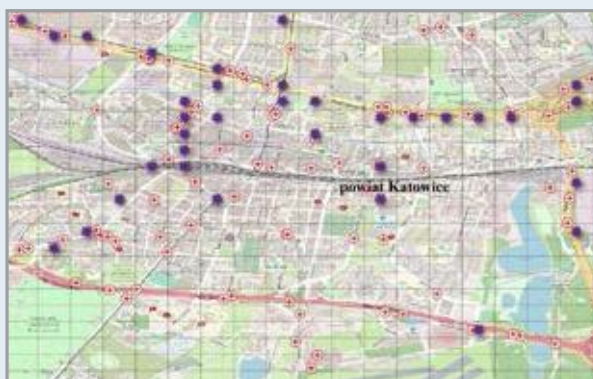
### Przetwarzanie danych

Nazwa miejscowości, ulica i numer drogi to zmienne kategoryczne. W procesie tworzenia nowych cech należy zamienić je na wartości numeryczne. Proces ten pozwoli na poprawną pracę modelu, którego można użyć do prognozowania.

Tak pobrane i przetworzone dane wykorzystano do budowy macierzy (siatki predykcji), tj. województwo śląskie podzielono

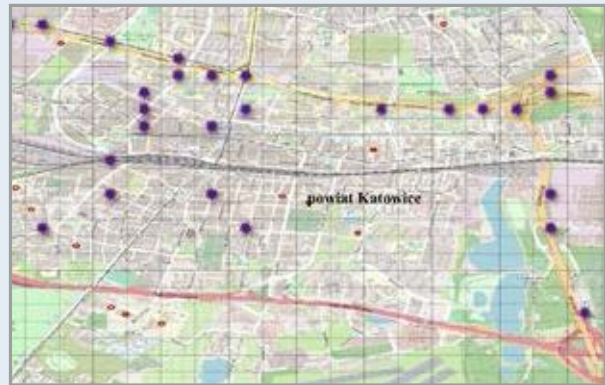
**Tabela 1.** Przykład struktury danych pobranych z SEWiK.

Lp.	GPS x	GPS y	Miejscowość	Ulica	Numer drogi	Liczba wypadków	Liczba zabitych w zdarzeniu	Liczba rannych w zdarzeniu
86891	19°11'524	50°17'328	Sosnowiec	Lenartowicza	P7609S	1	0	1
100890	18°54'325	50°10'449	Mikołów	Katowicka	G380085S	0	0	0
148722	19°00'589	50°08'005	Tychy	Dzwonkowa	P8003S	0	0	0
160215	19°15'239	50°38'038	Żarki-letnisko	Prosta	791	0	0	0
61219	19°07'471	50°49'041	Częstochowa	Warszawska	91	0	0	0

**Rys. 1.** Miasto Bytom z naniesioną siatką współrzędnych (przybliżenie obszaru całego województwa śląskiego).**Rys. 2.** Prognozowanie (test) z dokładnością 68% wypadków w 2018 r. (piktogram gwiazda: miejsce wytypowane, piktogram krzyż: rzeczywiste miejsce wypadku drogowego w 2018 r.).**Rys. 3.** Prognozowanie (test) z dokładnością 56% wypadków w 2018 r. (piktogram gwiazda: miejsce wytypowane, piktogram krzyż: rzeczywiste miejsce wypadku drogowego w 2018 r.).



## PRZEWIDYWANIE WYPADKÓW ZA POMOCĄ SZTUCZNEJ INTELIGENCJI

**Rys. 4.** Przewidywanie wypadków drogowych w 2019 r. z prawdopodobieństwem 68%.**Rys. 5.** Przewidywanie wypadków drogowych w 2019 r. z prawdopodobieństwem 56%.

na siatkę 898 561 prostokątów, z których każdy ma szerokość 10 sekund (długość geograficzna) i wysokość 5 sekund (szerokość geograficzna). W każdym polu jest zamieszczona informacja o wszystkich zdarzeniach zaistniałych w tym obszarze.

**Tworzenie nowych cech**

Po przetworzeniu danych kolejnym krokiem jest tworzenie nowych zmiennych – cech. W niniejszym projekcie dodano 3 nowe cechy, tj. średnią liczbę wypadków drogowych w danej miejscowości, na danej ulicy i drodze.

**Uczenie**

W niniejszym projekcie wykorzystano uczenie nadzorowane. Proces ten polegał na tym, iż na początku model otrzymał dane treningowe – bazę danych zawierających informacje o zdarzeniach drogowych z 2016 r. i jako „cel” ów model otrzymał dane dotyczące tylko miejsc, gdzie w 2017 r. zaistniały wypadki drogowe. Następnie taki proces poddano testom polegającym na tym, iż wprowadzono wszystkie dane o zdarzeniach drogowych zaistniałych w 2017 r. i przeprowadzono prognozowanie na 2018 r. Następnie porównano otrzymane wyniki z rzeczywistymi danymi dotyczącymi miejsc, gdzie zaistniały wypadki drogowe w 2018 r.

**Wizualizacja**

Dobierając odpowiednie parametry, uzyskano 2 rodzaje wyników:

- przy wytypowaniu 77 miejsc – precyzja typowania wyniosła 68%,
- przy wytypowaniu 261 miejsc – precyzja typowania wyniosła 56%.

Liczba wytypowanych miejsc odnosi się do całego województwa śląskiego. Poniżej przedstawiono przykłady wytypowanych miejsc w mieście Katowice.

Na rysunkach 2 i 3 przedstawiono wizualizację prognozowania miejsc wypadków drogowych w 2018 r. na podstawie danych z 2017 r. Natomiast na rysunkach 4 i 5 zaprezentowano wizualizację miejsc, w których może dojść do wypadku drogowego w 2019 r. na podstawie danych z 2018 r. Pod linkiem: [bit.ly/Wypadki2019](http://bit.ly/Wypadki2019) zamieszczono na mapie Google miejsca, w których według prognozy może dojść w 2019 r. do wypadku drogowego (kolor niebieski – prawdopodobieństwo 56%, czerwony – prawdopodobieństwo 68%).

**Wnioski**

Po przeprowadzaniu ww. kroków uzyskane wyniki porównano z tradycyjną metodą analizy danych. Porównując miejsca,

gdzie doszło do wypadku drogowego w 2017 r., z miejscami, gdzie doszło do wypadku drogowego w 2018 r., uzyskano dokładność 18%! Natomiast najniższy wynik, jaki uzyskano przy użyciu uczenia maszynowego, wyniósł 56%.

Na podstawie już uzyskanych wyników można poddać analizie wytypowane miejsca (dlaczego statystycznie prawdopodobieństwo wypadku drogowego jest tak duże) i dążyć do wyeliminowania zagrożenia.

Jest to początek drogi wykorzystania uczenia maszynowego do prognozowania miejsc, w których może dojść do wypadku drogowego. Baza SEWiK zawiera ogromną liczbę wysokiej jakości informacji na temat zdarzeń drogowych. Dodawanie cech do modelu, wykorzystanie już istniejących w celu budowania nowych może w znaczący sposób podnieść precyzję typowania. Następnie można przejść do zastosowania sztucznych sieci neuronowych lub wykorzystać dane np. z terenu całej Polski. Modele uczenia maszynowego działają znacznie lepiej, z wykorzystaniem bogatej „bazy wiedzy”. Próbując wykorzystać przedmiotowe techniki, najczęściej natrafia się na problem braku danych, natomiast w tym przypadku nie ma o tym mowy.

Jak widać na powyższym przykładzie, zastosowanie uczenia maszynowego może wspomóc pracę policjantów ruchu drogowego i być dodatkowym narzędziem w dążeniu do wizji ZERO wypadków ze skutkiem śmiertelnym. Jednocześnie należy podkreślić, iż uczenie maszynowe można zastosować również w takich dziedzinach, jak:

- prognozowanie występowania przestępstw,
- analiza sentymentu tekstu (wyszukiwanie działalności przestępczej w internecie), np. uprawa, produkcja lub handel narkotykami, rozpowszechnianie treści pedofilskich,
- analiza obrazu – kamery miejskie, m.in. poszukiwanie nietypowych zachowań ludzi, np. bójek, rozbojów.

**Summary*****Prevision of transportation accidents risk by using machine learning (artificial intelligence)***

In this article machine learning was presented to prevision regions, where traffic accidents may occur with injured or dead participants. Such prognosis can be achieved during so-called supervised learning, in which a model learns based on prior periods data. Thanks to applying this system it is possible to predict locations of accidents in time to come.

*Tłumaczenie: Jarosław Michalak*