

Uniwersytet Przyrodniczy we Wrocławiu

Wydział Biologii i Hodowli Zwierząt

Kierunek: Bioinformatyka

Studia stacjonarne drugiego stopnia

Bartosz Beżyk

117375

**Analiza porównawcza modeli uczenia  
maszynowego do przewidywania ciężkości  
obrażeń w wypadkach górskich**

Comparative analysis of machine learning models for predicting  
injury severity in mountain accidents

Praca wykonywana pod kierunkiem  
Dr. hab. inż. Heliodora Wierzbickiego  
Katedra Genetyki

Wrocław, 2024

### **Oświadczenie opiekuna pracy**

Oświadczam, że niniejsza praca została przygotowana pod moim kierownictwem i stwierdzam, że spełnia ona warunki do przedstawienia jej w postępowaniu o nadanie tytułu zawodowego.

Podpis opiekuna pracy:

Data:

### **Oświadczenie autora pracy**

Oświadczam, że niniejsza praca dyplomowa została napisana przeze mnie samodzielnie i nie zawiera treści uzyskanych w sposób niezgodny z obowiązującymi przepisami ani też nie była wcześniej przedmiotem procedur związanych z uzyskaniem tytułu zawodowego w wyższej uczelni.

Podpis autora pracy:

Data:

## Streszczenie

Wzrost popularności turystyki górskiej wiąże się ze zwiększoną liczbą wypadków, co stanowi wyzwanie dla służb ratowniczych. Prezentowana praca koncentruje się na analizie porównawczej modeli uczenia maszynowego w kontekście przewidywania ciężkości obrażeń w wypadkach górskich. Głównym celem pracy było zaprojektowanie i optymalizacja narzędzia analitycznego umożliwiającego dokładne prognozowanie ciężkości obrażeń poszkodowanych.

W badaniach wykorzystano dane zebrane przez GOPR w latach 2020-2024, obejmujące informacje o 9578 zdarzeniach. Praca obejmowała wstępne przetwarzanie i analizę danych, identyfikację istotnych zmiennych wpływających na ciężkość obrażeń oraz budowę i ocenę różnych modeli klasyfikacyjnych. Zastosowano takie metody jak korelacja rang Spearmana, SMOTE, RFECV, TOPSIS oraz modele Regresji Logistycznej, Drzew Decyzyjnych, Lasów Losowych, Drzew Wzmacnianych, Maszyn Wektorów Nośnych i klasyfikatory głosujące typu twardego oraz miękkiego. Modele porównano na podstawie obliczonych metryk takich jak dokładność, czułość, precyzja, miara F1, krzywa ROC oraz AUC.

Wyniki analizy wykazały, że model Drzew Wzmacnianych charakteryzował się najwyższą dokładnością (67%) w przewidywaniu ciężkości obrażeń, ze szczególnym uwzględnieniem wypadków śmiertelnych. Przeprowadzona analiza istotnych zmiennych wykazała, że czynniki takie jak rodzaj wypadku, przyczyna wypadku oraz czas trwania akcji mają największy wpływ na ciężkość obrażeń.

Praca dostarcza istotnych wniosków dotyczących zastosowania modeli uczenia maszynowego w ratownictwie górskim, a także wskazuje na kluczowe zmienne, które powinny być brane pod uwagę w procesie decyzyjnym. Uzyskane wyniki pozwalają na wystosowanie rekomendacji mających na celu usprawnienie pozyskiwania danych, ułatwienie oceny ciężkości obrażeń przez ratowników, a także formułowanie bardziej precyzyjnych strategii prewencyjnych i polityk publicznych związanych z bezpieczeństwem w górach.

## **Abstract**

The rising popularity of mountain tourism is associated with an increased number of accidents, posing a challenge for rescue services. This study focuses on a comparative analysis of machine learning models in the context of predicting injury severity in mountain accidents. The primary aim of the research was to design and optimize an analytical tool capable of accurately forecasting the severity of injuries sustained by victims.

The study utilized data collected by the Polish Mountain Search and Rescue (GOPR) from 2020 to 2024, comprising information on 9578 incidents. The study included preliminary data processing and analysis, identification of significant variables influencing injury severity, and the construction and evaluation of various classification models. Methods such as Spearman rank correlation, SMOTE, RFECV, TOPSIS, Logistic Regression, Decision Trees, Random Forests, Gradient Boosting Trees, Support Vector Machines, and hard and soft voting classifiers were employed. The models were compared based on indicators such as accuracy, sensitivity, precision, F1 score, ROC curve, and AUC.

The analysis demonstrated that the Gradient Boosting Trees model exhibited the highest accuracy (67%) in predicting injury severity, with particular emphasis on fatal accidents. The analysis of significant variables revealed that factors such as the type of accident, cause of the accident, and duration of the rescue operation have the greatest impact on injury severity.

This study provides valuable insights into the application of machine learning models in mountain rescue operations and highlights key variables that should be considered in the decision-making process. The findings enable the formulation of recommendations aimed at improving data collection, facilitating the assessment of injury severity by rescuers, and developing more precise preventive strategies and public policies related to mountain safety.

## Spis treści

1.	Wprowadzenie . . . . .	5
1.1.	Argumenty za prowadzeniem badań i zbieraniem danych . . . . .	5
1.2.	Przewidywanie ciężkości obrażeń poszkodowanych . . . . .	5
1.3.	GOPR i wypadki górskie . . . . .	6
1.4.	Modele predykcyjne i klasyfikacyjne . . . . .	7
1.5.	Jakość danych . . . . .	8
1.6.	Detekcja istotnych czynników . . . . .	8
2.	Cel pracy . . . . .	10
3.	Materiały i Metody . . . . .	11
3.1.	Materiały . . . . .	11
3.2.	Metody . . . . .	17
3.2.1.	Kodowanie jednoelementowe . . . . .	17
3.2.2.	Korelacja rang Spearmana . . . . .	17
3.2.3.	SMOTE . . . . .	18
3.2.4.	RFECV . . . . .	19
3.2.5.	Model Regresji Logistycznej . . . . .	20
3.2.6.	Model Drzew Decyzyjnych . . . . .	21
3.2.7.	Model Lasów Losowych . . . . .	22
3.2.8.	Model Drzew Wzmacnianych . . . . .	23
3.2.9.	Model Maszyny Wektorów Nośnych . . . . .	24
3.2.10.	Klasyfikatory głosujące . . . . .	25
3.2.11.	Istotność zmiennych . . . . .	26
3.2.12.	Metryki dokładności . . . . .	27
3.2.13.	TOPSIS . . . . .	28
4.	Wyniki . . . . .	31
5.	Dyskusja . . . . .	52
6.	Podsumowanie i wnioski . . . . .	56

# 1. Wprowadzenie

Wraz z rosnącą popularnością turystyki górskiej, wzrasta liczba wypadków w trudno dostępnych terenach, co stanowi wyzwanie dla służb ratowniczych. Opracowanie skutecznych metod przewidywania ciężkości obrażeń u poszkodowanych może znacząco wpłynąć na efektywność i szybkość reakcji, a także na optymalizację procesu ratowniczego i późniejszego leczenia.

## 1.1. Argumenty za prowadzeniem badań i zbieraniem danych

Prowadzenie badań nad wypadkowością oraz zbieranie danych wysokiej jakości na temat wypadków w górach jest niezbędne z wielu względów. Nie tylko pomaga to w tworzeniu efektywnych strategii prewencyjnych, ale również znacząco wpływa na zwiększenie bezpieczeństwa turystów oraz ratowników górskich. Przesłankami przemawiającymi za wymienionymi działaniami są:

- **Udoskonalenie metod ratownictwa:** zrozumienie przyczyn wypadków w górach jest kluczowe dla opracowania skutecznych technik ratowniczych, które mogą być specyficzne np. dla różnych rodzajów terenu czy warunków pogodowych. Przykładem może być badanie przyczyn lawin, gdzie szczegółowe dane mogą pomóc w przewidywaniu ich występowania i ostrzeganiu turystów oraz ratowników przed potencjalnym zagrożeniem [1].
- **Poprawa strategii prewencyjnych:** efektywne strategie prewencyjne oparte na dokładnych danych mogą znacznie obniżyć liczbę wypadków w górach. Przykładem może być rozwój aplikacji lub systemów ostrzegawczych wykorzystujących dane o pogodzie, ruchu turystycznym czy historii wypadków na danej trasie [2].
- **Zwiększenie bezpieczeństwa turystów:** dokładne i łatwo dostępne informacje o potencjalnych zagrożeniach oraz optymalnych ścieżkach w górach mogą znacząco zwiększyć bezpieczeństwo osób poruszających się po górskich terenach. Informowanie o trudnościach tras, przewidywanym czasie przejścia, a także aktualnych warunkach może pomóc turystom lepiej przygotować się do wycieczek [3].
- **Rozwój polityk publicznych i ustawodawstwa:** dokładne dane mogą być wykorzystywane przez władze lokalne i narodowe do tworzenia lub modyfikacji przepisów dotyczących bezpieczeństwa w górach. Może to obejmować zarówno regulacje dotyczące wyposażenia, jak i zasad poruszania się po określonych obszarach [4].

## 1.2. Przewidywanie ciężkości obrażeń poszkodowanych

### Definicja problemu

Modelowanie predykcyjne w kontekście ratownictwa górskiego może obejmować stosowanie algorytmów uczenia maszynowego lub modelowania statystycznego do przewidywania ciężkości obrażeń na podstawie danych historycznych, które mogą obejmować informacje o

typie wypadku, warunkach atmosferycznych, charakterystyce terenu oraz o samych poszkodowanych (np. wiek, płeć, istniejące schorzenia). Potencjalnymi korzyściami z modelowania predykcyjnego są:

- **Lepsze zrozumienie okoliczności:** może być głównym czynnikiem wspomagającym podejmowanie środków prewencyjnych w celu zminimalizowania ilości i powagi wypadków w górach [5].
- **Zwiększenie szybkości reakcji:** szybka identyfikacja ciężkości obrażeń może skrócić czas potrzebny na organizację odpowiedniej pomocy medycznej.
- **Optymalizacja zasobów:** efektywniejsze zarządzanie dostępnymi zasobami medycznymi i ratowniczymi.
- **Poprawa wyników leczenia:** im szybsza i lepiej dostosowana pomoc, tym lepsze prognozy dla poszkodowanych.
- **Wsparcie decyzji:** model predykcyjny może wspierać decyzje ratowników o priorytecie działań i ewakuacji, szczególnie w skomplikowanych lub wielu wypadkach jednocześnie.

### 1.3. GOPR i wypadki górskie

”Górskie Ochotnicze Pogotowie Ratunkowe (GOPR) jest stowarzyszeniem kultury fizycznej o zasięgu ogólnokrajowym. Terenem działania jest obszar Rzeczypospolitej Polskiej ze szczególnym uwzględnieniem regionów górskich w kraju i poza jego granicami. GOPR działa na podstawie ustawy o bezpieczeństwie i ratownictwie w górach i na zorganizowanych terenach narciarskich z dnia 18 sierpnia 2011 roku [6]. GOPR swoje zadania statutowe – niesienie pomocy w górach ludziom, których zdrowie lub życie jest zagrożone, zapobieganie wypadkom w górach oraz ochrona środowiska górskiego – realizuje poprzez 7 Grup Regionalnych, tj:

- Grupę Beskidzką
- Grupę Bieszczadzką
- Grupę Jurajską
- Grupę Karkonoską
- Grupę Krynicką
- Grupę Podhalańską
- Grupę Sudecką

Obszar jaki zabezpiecza GOPR to pasma gór polskich od Karkonoszy poprzez Jurę Krakowsko-Częstochowską do Bieszczad, o łącznej powierzchni 20.410 km<sup>2</sup> [7]”.

Turystyka górska w Polsce zyskuje na popularności z roku na rok. Coraz więcej osób uprawia takie formy aktywności jak kolarstwo górskie, narciarstwo skiturowe czy biegi górskie. Wraz ze wzrostem liczby osób poruszających po górach, wzrasta liczba wypadków z tym związanych [8], [9].

Za ratownictwo górskie w Polsce odpowiadają dwie organizacje:

- GOPR - Górskie Ochotnicze Pogotowie Ratunkowe
- TOPR - Tatrzańskie Ochotnicze Pogotowie Ratunkowe

Ratownicy górscy stoją przed wyzwaniem udzielania pomocy w trudnych technicznie i skomplikowanych operacjach. Osoby potrzebujące pomocy często znajdują się w trudno dostępnych terenach, oddalonych wiele kilometrów od najbliższej stacji ratunkowej. Dotarcie do poszkodowanych może być długotrwałe, trudne i skomplikowane logistycznie. Udzielenie pomocy poszkodowanemu to często nie jest ostatnie zadanie, które ratownicy muszą wykonać, gdyż zwykle czeka ich jeszcze ewakuacja rannego. Podczas tej czynności poszkodowany jest stale monitorowany, a w odpowiedzi na jego stan podejmowane są zarówno czynności medyczne, jak i logistyczne.

Bardzo ważnym aspektem podczas prowadzenia działań jest prawidłowa ocena stanu, w którym znajduje się poszkodowany, gdyż to na jej podstawie podejmowane są decyzje o rodzaju i priorytecie udzielanej pomocy.

#### **1.4. Modele predykcyjne i klasyfikacyjne**

Wykorzystanie modeli predykcyjnych i klasyfikacyjnych stało się nieodzownym narzędziem w nowoczesnej medycynie i ratownictwie. Umożliwiają one precyzyjne prognozowanie wyników leczenia, ocenę ryzyka wystąpienia określonych stanów zdrowotnych oraz optymalizację procesów decyzyjnych w nagłych przypadkach. Dzięki temu możliwe jest lepsze planowanie działań terapeutycznych i ratunkowych, co przyczynia się do poprawy jakości opieki zdrowotnej i skuteczności interwencji ratunkowych [10].

Modele predykcyjne są używane do przewidywania ciągłych wartości lub zmiennych. Zastosowania tych modeli obejmują prognozowanie cen akcji, przewidywanie popytu na produkty, szacowanie wartości parametrów życiowych czy estymację wartości nieruchomości. Modele te starają się przewidzieć dokładną wartość na podstawie dostępnych danych wejściowych [11].

Modele klasyfikacyjne służą do przewidywania kategorii lub klasy, do której należy dany obiekt. Zamiast przewidywać konkretną wartość numeryczną, model klasyfikacyjny określa, do której z predefiniowanych kategorii należy dana obserwacja. Na przykład, model klasyfikacyjny może być używany do oceny ciężkości obrażeń w wypadkach drogowych lub do diagnozowania chorób na podstawie wyników badań medycznych [12], [13].

Algorytmy klasyfikacyjne, są szeroko używane w różnych dziedzinach nauki, takich jak medycyna [14], meteorologia [15] czy statystyka [16]. Podobne podejście stosowane było w analizach wypadków drogowych. Z uwagi na podobną specyfikę przewidywań zmiennych, doświadczenia zebrane z zastosowań tych algorytmów do analizy wypadków drogowych, mogą służyć jako wskazówki do analizy wypadków górskich. Przykładowo, w badaniach przedstawionych w pracach [1], [17], [18], [19] wykorzystano zaawansowane metody statystyczne oraz algorytmy uczenia maszynowego, takie jak Lasy Losowe i głębokie sieci neuronowe do analizy i przewidywania ryzyka wypadku, czy rozproszenia uwagi kierowcy. Ponadto, użyte zostały



do oszacowania efektów leczenia, co ma kluczowe znaczenie w analizie bezpieczeństwa ruchu drogowego. Dzięki tym metodom możliwe jest znaczące zwiększenie precyzji prognoz, co przekłada się na lepszą efektywność interwencji ratunkowych oraz projektów infrastrukturalnych.

### 1.5. Jakość danych

Skuteczność modeli predykcyjnych zależy nie tylko od zastosowanych algorytmów, lecz w dużej mierze od dokładności i kompletności danych o wypadkach. W przypadku modelowania ciężkości obrażeń, dane wysokiej jakości umożliwiają lepsze zrozumienie czynników ryzyka i skuteczniejsze planowanie działań prewencyjnych. Nieprecyzyjne dane mogą prowadzić do niedoszacowania ryzyka, co zagraża skuteczności interwencji [20], [21], [22], [23].

### 1.6. Detekcja istotnych czynników

Znalezienie istotnych przyczyn wpływających na jakość predykcji obrażeń w wypadkach jest kluczowe zarówno z punktu widzenia modelowania statystycznego, jak i praktycznej prewencji.

#### **Dlaczego jest to ważne:**

- **Zwiększenie dokładności modeli:** modele predykcyjne, takie jak Lasy Losowe, wykorzystują różnorodne zmienne do przewidywania ciężkości obrażeń w wypadkach. Poprzez identyfikację i zastosowanie najbardziej wpływowych czynników (takich jak warunki pogodowe, typ drogi, prędkość pojazdów) można znacząco zwiększyć dokładność tych predykcji. Efektywne modele przewidujące pozwalają na lepsze zrozumienie zależności i przyczyn ciężkich wypadków, co jest kluczowe dla zwiększenia ich skuteczności i wiarygodności [24].
- **Optymalizacja zasobów:** dokładne modele predykcyjne umożliwiają efektywniejsze zarządzanie zasobami w sytuacjach kryzysowych. Na przykład, dokładna predykcja ciężkości obrażeń pozwala służbom ratunkowym na lepsze planowanie odpowiedzi, co może obejmować alokację ambulansów, straży pożarnej i policji w miejscach, gdzie są najbardziej potrzebne [4].
- **Prewencja wypadków:** poprzez analizę czynników przyczyniających się do ciężkich obrażeń, można wprowadzić skuteczne strategie zapobiegawcze. Na przykład, jeśli analiza wskaże, że wiele poważnych wypadków ma miejsce na skutek nadmiernej prędkości na mokrych nawierzchniach, lokalne władze mogą zastosować ograniczenia prędkości obowiązujące w trakcie deszczu lub poprawić drenaż drogowy [1].
- **Ustawodawstwo i polityka publiczna:** skuteczne modele mogą również wpłynąć na kształtowanie polityki publicznej i przepisów drogowych. Jeśli na przykład okaże się, że niektóre typy pojazdów częściej uczestniczą w wypadkach skutkujących ciężkimi obrażeniami, możliwe będzie wprowadzenie dodatkowych przepisów regulujących ich użyt-

kowanie [2].

- **Edukacja i świadomość społeczna:** zrozumienie czynników wpływających na ciężkość wypadków pozwala na lepsze informowanie i edukowanie społeczeństwa o ryzyku i środkach zapobiegawczych, co może obniżyć liczbę wypadków i ich ciężkość [3].

**Korzyści:**

- **Zmniejszenie liczby wypadków i ich ciężkości:** efektywne modele predykcyjne i wynikające z nich działania prewencyjne mogą znacząco zmniejszyć liczbę wypadków oraz ich ciężkość, co bezpośrednio przekłada się na obniżenie kosztów zdrowotnych i humanitarnych.
- **Zmniejszenie kosztów gospodarczych:** wypadki generują duże koszty związane z opieką zdrowotną, utratą produktywności, a także zniszczeniem mienia. Działania wynikające z dokładnych modeli predykcyjnych mogą ograniczyć te straty.
- **Podniesienie bezpieczeństwa publicznego:** zmniejszenie liczby wypadków i ich ciężkości bezpośrednio wpływa na poprawę bezpieczeństwa publicznego, co jest fundamentalnym celem zarządzania kryzysowego i polityki bezpieczeństwa wewnętrznego oraz jednym z punktów statutowych Górskiego Ochotniczego Pogotowia Ratunkowego [25].

## 2. Cel pracy

Celem niniejszej pracy jest zaprojektowanie i wybór zaawansowanego narzędzia analitycznego, które będzie służyć do przewidywania ciężkości obrażeń wynikających z wypadków górskich. Główne założenia pracy obejmują:

- **Optymalizację metod pozyskiwania danych**
- **Identyfikację i analizę czynników wpływających na obrażenia**
- **Stworzenie modelu klasyfikacyjnego**
- **Walidację i weryfikację modelu**

Praca ta przyczyni się do lepszego zrozumienia mechanizmów i czynników wpływających na ciężkość wypadków na terenach górskich, co ma kluczowe znaczenie dla poprawy bezpieczeństwa i efektywności działań ratowniczych. Ponadto, wyniki te mogą być wykorzystane do formułowania rekomendacji dla organizacji zajmujących się bezpieczeństwem w górach, a także do rozwijania polityk publicznych związanych z turystyką górską i ratownictwem.

### 3. Materiały i Metody

#### 3.1. Materiały

Dane wykorzystane w pracy pochodzą z kart wypadków górskich pochodzących z siedmiu grup regionalnych GOPR, zebranych w latach 2020-2024 (stan na 05.03.2024). Zawierają informacje o 9578 zdarzeniach, które miały miejsce na terenie działania Górskiego Ochotniczego Pogotowia Ratunkowego. Każdy rekord zawierał informację o:

- grupie regionalnej, w której wydarzył się wypadek,
- godzinie zgłoszenia,
- miesiącu zdarzenia,
- fazie tygodnia (koniec tygodnia lub dni robocze),
- lokalizacji zdarzenia,
- miejscu dyżuru ratowników zadysponowanych do działania,
- liczbie ratowników biorących udział w akcji,
- przyczynie wypadku,
- rodzaju wypadku,
- wysokości miejsca wypadku (czy zdarzenie miało miejsce powyżej 1500 m.n.p.m.),
- czy ratownicy zostali zadysponowani lub sami kontaktowali się z Centrum Powiadamiania Ratunkowego,
- wieku poszkodowanego,
- rodzaju transportu/ewakuacji poszkodowanego z miejsca zdarzenia,
- pogodzie,
- płci poszkodowanego,
- czasie trwania działań,
- obrażeniach poszkodowanego.

Powyższe informacje, zwane dalej zmiennymi (ang. variables), przyjmowały wartości numeryczne lub zostały przypisane do poszczególnych kategorii. Poniżej znajduje się ich dokładny opis.

#### **Grupa**

Zmienna zawiera informacje o grupie, na terenie działania której doszło do zdarzenia. Przyjmuje siedem wartości:

Tabela 1: Grupy, na terenie działania których doszło do zdarzenia

Numer	Grupa
1	Sudecka
2	Beskidzka
3	Bieszczadzka
4	Karkonoska
5	Krynicka
6	Jurajska

Tabela 1 – Kontynuacja z poprzedniej strony

Numer	Grupa
7	Podhalańska

### Godzina

Zmienna zawiera informację o godzinie zgłoszenia wypadku do ratowników GOPR. Przyjmuje wartości liczbowe od 1 do 24, odpowiadające każdej godzinie doby, w której mogło nastąpić zgłoszenie.

### Miesiąc

Zmienna informuje o miesiącu, w którym zdarzył się wypadek. Przyjmuje wartości od 1 do 12, odpowiadające poszczególnym miesiącom.

### Faza tygodnia

Zmienna została utworzona przez podział tygodnia na dni robocze i dni wolne od pracy. Przyjmuje ona dwie wartości:

Tabela 2: Podział tygodnia na dni robocze i dni wolne od pracy

Wartość	Opis
1	Dni od poniedziałku do czwartku (dni robocze)
2	Dni od piątku do niedzieli (dni wolne od pracy)

Rozważano alternatywne metody podziału tygodnia, w tym podział na poszczególne dni lub na grupę od poniedziałku do piątku oraz sobota-niedziela. Wybrana metoda podziału, przedstawiona powyżej, okazała się najbardziej efektywna w kontekście wydajności tworzonych modeli.

### Lokalizacja

Zmienna zawiera informacje o lokalizacji wypadku. Nie odnoszą się one do dokładnych współrzędnych miejsca zdarzenia, lecz do przypisanych nazw pasm lub masywów górskich. Obejmuje ona 25 kategorii, które zostały przedstawione poniżej:

Tabela 3: Lokalizacje wypadków

Numer	Lokalizacja
1	Karkonosze
2	Masyw Śnieżnika
3	Góry Orlickie
4	Góry Sowie
5	Góry Stołowe
6	Pasma Baraniej Góry
7	Pieniny
8	Góry Izerskie
9	Pasma Barniej Góry
10	Pasma Pilska
11	Masyw Babiej Góry
12	Pasma Czantorii
13	Gorce

Kontynuacja na następnej stronie

**Tabela 3 – kontynuacja z poprzedniej strony**

Numer	Lokalizacja
14	Beskid Wyspowy
15	Beskid Mały
16	Góry Wałbrzyskie
17	Beskid Śląski
18	Inne
19	Beskid Żywiecki
20	Żar
21	Rudawy Janowickie
22	Beskid Sądecki
23	Beskid Niski
24	Bieszczady
25	Jura Krakowsko-Częstochowska

## Dyżur

Zmienna zawiera informację o miejscu dyżuru ratowników GOPR, których zadysponowano do wypadku. Zawiera 52 kategorie przedstawione poniżej:

**Tabela 4: Unikalne wartości zmiennej 'Dyżur'**

Numer	Miejsce dyżuru
1	Stacja Karpacz
2	Międzygórze
3	Zieleniec
4	Sokolec
5	Stacja Szklarska Poręba
6	Beskid Śląski
7	Szczawnica
8	CSR Rabka Zdrój - DZIEŃ, Szczawnica
9	CSR Szczyrk - DZIEŃ
10	Pasmo Piłska
11	Masyw Babiej Góry
12	Turbacz Dyżurka
13	Centrala - Jelenia Góra - DZIEŃ- Karpacz
14	Centrala - Jelenia Góra - NOC / KARPACZ
15	CSR Rabka Zdrój - NOC
16	Karlów
17	CSR Szczyrk - NOC
18	Stacja Centralna
19	CSR Rabka Zdrój - DZIEŃ
20	CSR Rabka Zdrój - DZIEŃ, Turbacz Dyżurka
21	Ochotnica
22	CSR Szczyrk - DZIEŃ, Hala Skrzyczneńska
23	Krościenko
24	Beskid Mały
25	Przysłop
26	Stacja Świeradów Zdrój
27	Centrala - Jelenia Góra - DZIEŃ- Karpacz, Stacja Szklarska Poręba
28	Śnieżnica schronisko
29	Nowy Targ - Długa Polana
30	CSR Krynica - DZIEŃ
31	TPR Banica
32	RSR Piwniczna
33	TSR Krempna - MPN

Tabela 4 – Kontynuacja z poprzedniej strony

Numer	Nazwa dyżuru
34	Dyspozycja
35	Centrala - Jelenia Góra - NOC / KARPACZ, Dyspozycja
36	TSR Jaworzyna
37	TSR Prehyba
38	CSR Krynica-NOC
39	zabezpieczenie zawodów sportowych
40	Krościenko, Szczawnica
41	Nowy Targ - Długa Polana, Turbacz Dyżurka
42	ŻAR - LATO
43	WISŁA
44	RSR Cisna - DZIEŃ
45	RSR Ustrzyki Górne - DZIEŃ
46	Stacja Centralna Sanok - DZIEŃ
47	DS Dukla - DZIEŃ
48	RSR Ustrzyki Górne - NOC
49	Stacja Centralna Sanok - NOC
50	Stacja Centralna - Dzień
51	Dolinki - Dzień
52	Olsztyn - Dzień

### Liczba ratowników

Zmienna zawiera informacje o liczbie ratowników zadysponowanych do działania. Przyjmuje wartości liczbowe, całkowite od 0 (np. podczas telefonicznego naprowadzenia osoby zagubionej na szlak) do 68.

### Przyczyna wypadku

Zmienna informuje o przyczynie wypadku. Przyjmuje jedenaście wartości opisanych poniżej:

Tabela 5: Rodzaje przyczyn zdarzeń

Numer	Przyczyna
1	Upadek
2	Upadek z wysokości
3	Zabłądzenie
4	Brak wiedzy/sprzętu
5	Lawina
6	Niska temperatura
7	Wysoka temperatura
8	Ukąszenie/ugryzienie
9	Zachorowanie
10	Inne
11	Brak obiektywnej przyczyny

### Rodzaj wypadku

Zmienna opisuje do jakiej kategorii wypadków zaliczono zdarzenie. Przyjmuje dziewięć wartości:

Tabela 6: Rodzaje wypadków

Numer	Rodzaj wypadku
1	Narciarski
2	Turystyczny
3	Ekstremalny
4	Rowerowy
5	Trening sportowy
6	Prace leśne
7	Zbędny alarm
8	Samobójstwo
9	Inne

**Uwagi:**

Kategoria wypadku *Ekstremalny* została utworzona poprzez zakwalifikowanie mniej licznych, ale istotnych kategorii, tj. wypadek wspinaczkowy, wypadek jaskiniowy, wypadek paralotniowy.

**Powyżej 1500 m.n.p.m.**

Zmienna informuje tym, czy wypadek miał miejsce poniżej czy powyżej wysokości 1500 m.n.p.m. Przyjmuje dwie wartości:

Tabela 7: Wysokość, na której doszło do wypadku

Kod wysokości	Opis wysokości
0	Poniżej 1500 m n.p.m.
1	Powyżej 1500 m n.p.m.

**CPR**

Zmienna ta określa, czy podczas wypadku nawiązano współpracę z Centrum Powiadomienia Ratunkowego. Może to oznaczać, że ratownicy zostali zadysponowani przez CPR lub że sami zgłosili potrzebę wsparcia, na przykład przez Zespół Ratownictwa Medycznego.

Tabela 8: Współpraca z Centrum Powiadomienia Ratunkowego

Kod	Status współpracy
0	Nie nawiązano współpracy podczas wypadku
1	Nawiązano współpracę podczas wypadku

**Wiek**

Zmienna informuje o wieku poszkodowanej osoby. Przyjmuje wartości liczbowe, całkowite od 1 do 99.

**Rodzaj transportu/ewakuacji**

Zmienna informuje o głównym sposobie transportu/ewakuacji poszkodowanego z miejsca zdarzenia. Przyjmuje pięć wartości:



Tabela 9: Rodzaje transportu/ewakuacji poszkodowanego

Numer	Rodzaj transportu/ewakuacji
1	Ręczny
2	Pojazd
3	Śmigłowiec
4	Nie dotyczy
5	Inne

## Pogoda

Zmienna opisuje stan pogody podczas wypadku. Przyjmuje dwie wartości:

Tabela 10: Opis pogody podczas wypadku

Wartość	Stan pogody
1	Dobra
2	Zła

### Uwagi:

Pogoda została uznana za złą, gdy wystąpił przynajmniej jeden z poniższych czynników: silny wiatr, opad deszczu/śniegu, ograniczona widoczność.

## Płeć

Zmienna zawiera dwie kategorie informujące o płci osoby poszkodowanej.

Tabela 11: Kategorie informujące o płci osoby poszkodowanej

Symbol	Płeć
k	Kobieta
m	Mężczyzna

## Czas akcji

Zmienna zawiera informacje o czasie trwania poszczególnych działań. Czas wyrażony jest w minutach i przyjmuje wartości od 0 do 600.

## NACA

Zmienna opisuje ciężkość obrażeń poszkodowanego. Została oceniona przez ratowników biorących udział w działaniach. W pracy rozróżniono cztery klasy przyjmujące wartości:

Tabela 12: Klasyfikacja ciężkości obrażeń

Wartość	Opis obrażeń
0	Brak obrażeń
1	Obrażenia lekkie
2	Obrażenia średnie i ciężkie
3	Zgon

Wyróżnione klasy są wynikiem agregacji kategorii skali NACA [26]. Podczas prac nad projektem rozważano różne metody podziału, w tym zachowanie oryginalnej skali, jednak za-

proponowane rozwiązanie okazało się najskuteczniejsze pod kątem wydajności opracowywanych modeli. Oryginalna skala NACA, używana przez ratowników, została przedstawiona w Tabeli poniżej:

Tabela 13: Oryginalna skala NACA i odpowiadające kategorie obrażeń

NACA	Opis obrażeń
I	Brak obrażeń
II	Powierzchnowe urazy
III	Wskazane leczenie ambulatoryjne
IV	Wskazane leczenie szpitalne
V	Potencjalne zagrożenie życia
VI	Zagrożenie życia
VII	Stan po skutecznej resuscytacji
VIII	Zgon

Klasyfikacja obrażeń jest szczególnie istotna w kontekście analizy wypadkowości. Wyodróżnienie klasy zgonu pozwala na dokładniejszą analizę przyczyn i okoliczności prowadzących do wypadków śmiertelnych. Rozpoznanie tych czynników może przyczynić się do lepszego zrozumienia i redukcji tych wypadków. Wprowadzenie tej klasy doprowadziło do stworzenia zestawu danych o następujących liczebnościach:

Tabela 14: Liczebności poszczególnych klas ciężkości obrażeń

Klasa ciężkości	Liczebność
0	2662
1	3105
2	3507
3	184

## 3.2. Metody

### 3.2.1. Kodowanie jednoelementowe

W ramach przetwarzania danych zastosowano technikę kodowania jednoelementowego, znaną również jako one-hot encoding. Jest to metoda konwersji zmiennych kategorycznych na format, który może być bezpośrednio używany w modelach uczenia maszynowego. Każda kategoria zmiennej jest transformowana w osobną kolumnę z wartościami binarnymi 0 lub 1, gdzie 1 oznacza obecność danej kategorii, a 0 jej brak [52].

Technika jest użyteczna w modelach, które nie radzą sobie dobrze z danymi kategorycznymi lub wymagają danych wejściowych w postaci numerycznej, jak regresja logistyczna. Kodowanie jednoelementowe eliminuje także problem fałszywej hierarchii numerycznej, który może pojawić się w przypadku kodowania zmiennych kategorycznych za pomocą pojedynczych liczb.

### 3.2.2. Korelacja rang Spearmana

W celu zbadania związku między różnymi rodzajami zmiennych użytych w badaniu, zastosowano metodę korelacji rang Spearmana, która efektywnie radzi sobie z danymi o niestandardowych rozkładach i nieliniowych zależnościach.

Korelacja rang Spearmana jest nieparametryczną metodą statystyczną wykorzystywaną do określania siły i kierunku związku między dwiema zmiennymi. Metoda ta nie zakłada normalności rozkładów zmiennych, co czyni ją użyteczną w analizie danych, które mogą nie spełniać założeń parametrycznych metod korelacji [27].

### **Proces rangowania**

Proces rangowania polega na przypisywaniu kolejnych liczb porządkowych wszystkim wartościom badanych zmiennych. Wartości są sortowane, a każdej z nich przypisywana jest pozycja w tej uporządkowanej sekwencji. W przypadku wartości powtarzających się, zastosowanie znajduje metoda przypisania średniej rangi, która równomiernie rozdziela rangi między powtarzające się wartości.

### **Obliczanie korelacji Spearmana**

Po przypisaniu rang, korelacja Spearmana między dwiema zmiennymi  $X$  i  $Y$ , którym przypisano rangi  $R_X$  oraz  $R_Y$ , jest obliczana przy użyciu następującego wzoru:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (1)$$

gdzie:

- $d_i = R_{X_i} - R_{Y_i}$  – różnica rang między odpowiadającymi sobie wartościami zmiennych  $X$  i  $Y$  dla  $i$ -tej obserwacji,
- $n$  – liczba obserwacji (par wartości) w analizowanych danych.

Wzór ten efektywnie mierzy stopień związku poprzez analizę kwadratów różnic rang. Suma kwadratów różnic jest następnie skalowana względem maksymalnej wartości, jaką może przyjąć dla danej liczby obserwacji, co pozwala na interpretację wyniku w standardowym zakresie od  $-1$  do  $1$ , gdzie wartości bliskie  $-1$  lub  $1$  wskazują na silny związek, a wartość  $0$  na brak związku.

### **3.2.3. SMOTE**

W celu zrównoważenia klas ciężkości obrazów w niezbalansowanym zestawie danych wykorzystanym w pracy, zastosowano technikę SMOTE (ang. Synthetic Minority Over-sampling Technique).

Technika ta została opracowana w celu poprawy klasyfikacji w niezrównoważonych zestawach danych, gdzie jedna z klas jest znacząco mniej liczna niż inne [28].

### **Mechanizm działania**

Metoda polega na syntetycznym generowaniu nowych przykładów klasy mniejszościowej. Nowe próbki są tworzone przez interpolację liniową pomiędzy istniejącymi próbkami tej

klasy.

SMOTE tworzy nowe instancje przez losowe wybieranie pary istniejących przykładów klasy mniejszościowej i generowanie nowej instancji jako ich kombinacja liniowa.

Matematycznie, generowanie nowej próbki  $x_n$  może być opisane wzorem:

$$x_n = x_i + \lambda \cdot (x_z - x_i) \quad (2)$$

gdzie:

- $x_i$  - wektor zmiennych jednej z losowo wybranych próbek klasy mniejszościowej,
- $x_z$  - wektor zmiennych drugiej, różnej od  $x_i$ , losowo wybranej próbki tej samej klasy mniejszościowej,
- $\lambda$  - losowo wybrana liczba z przedziału  $[0,1]$ , która determinuje, jak daleko wzdłuż odcinka łączącego  $x_i$  i  $x_z$  znajdzie się nowa próbka  $x_n$ .

Technika ta jest efektywnym rozwiązaniem problemów związanych z nie zrównoważonymi zestawami danych, umożliwiając klasyfikatorom osiągnięcie lepszej dokładności poprzez zrównoważenie klasy mniejszościowej w zestawie danych.

Do implementacji techniki SMOTE w języku Python użyty został moduł `imbalanced-learn` [29].

#### 3.2.4. RFECV

W pracy wykorzystano technikę RFECV (ang. Recursive Feature Elimination with Cross-Validation) do optymalizacji modeli poprzez eliminację najmniej istotnych zmiennych.

Technika RFECV to metoda selekcji zmiennych stosowana w uczeniu maszynowym, mająca na celu optymalizację modeli poprzez eliminację najmniej istotnych zmiennych z zestawu danych. Jest to istotne w zadaniach, gdzie wysoka wymiarowość danych może prowadzić do przetrenowania i obniżenia ogólnej zdolności generalizacji modelu [30].

##### **Mechanizm działania**

RFECV łączy technikę rekurencyjnego eliminowania zmiennych (RFE - ang. Recursive Feature Elimination)[30] z walidacją krzyżową [31], zapewniając bardziej wiarygodną ocenę wpływu poszczególnych zmiennych na wydajność modelu. Algorytm rozpoczyna pracę z pełnym zestawem zmiennych, a następnie iteracyjnie usuwa najmniej istotne zmienne, jednocześnie oceniając model po każdej eliminacji, aby stwierdzić, czy usunięcie zmiennych prowadzi do poprawy wydajności modelu.

##### **Walidacja krzyżowa**

Walidacja krzyżowa (ang. Cross validation) to technika oceny generalizacji modeli statystycznych poprzez podział danych na kilka podzbiorów. Model jest trenowany na większości tych podzbiorów, a testowany na pozostałych, co jest powtarzane wielokrotnie, aby zapewnić, że każdy podzbiór służył zarówno jako dane treningowe, jak i testowe. Ta metoda jest szeroko stosowana w uczeniu maszynowym do oceny wydajności modeli, zwłaszcza gdy dostępne są

ograniczone dane [32].

Ocena modelu poprzez walidację krzyżową w kontekście RFECV polega na obliczaniu miar wydajności, takich jak dokładność (ang. accuracy), w każdym z cykli walidacji. Umożliwia to obiektywne porównanie modeli z różnymi zestawami zmiennych i wybór tego, który zapewnia najlepszą wydajność na niezależnych podzbiorach danych [32].

### **Implementacja**

Technika RFECV została zaimplementowana przy użyciu biblioteki `scikit-learn` [33]. Jako klasyfikatora domyślnego użyto modelu *GradientBoostingClassifier*

### **3.2.5. Model Regresji Logistycznej**

W prezentowanych badaniach stworzono i przetestowano sześć modeli: model regresji logistycznej, model drzew decyzyjnych, model lasów losowych, model drzew wzmacianych, model maszyn wektorów nośnych oraz klasyfikatory głosujące.

Pierwszym stworzonym modelem był model regresji logistycznej.

Model regresji logistycznej to specyficzny przypadek uogólnionego modelu liniowego (GLM - ang. Generalized Linear Model), który pozwala na modelowanie zależności pomiędzy zmienną zależną a jednym lub więcej predyktorami, które są modelowane przy użyciu funkcji logitowej jako funkcji łączącej [34].

### **Funkcja logit**

Funkcja logitowa, używana do modelowania logarytmu szans, opisuje logarytm stosunku prawdopodobieństwa zajścia zdarzenia do prawdopodobieństwa jego niezajścia:

$$\text{logit}(p) = \log \left( \frac{p}{1-p} \right), \quad (3)$$

gdzie  $p$  jest prawdopodobieństwem przynależności do klasy zainteresowania.

Transformacja logitowa jest kluczowa, gdyż mapuje prawdopodobieństwo z przedziału (0,1) na całą oś rzeczywistą, co jest niezbędne do zastosowania liniowego modelowania statystycznego [35].

### **Implementacja**

Model implementowany w niniejszej analizie wykorzystuje funkcję `LogisticRegression` z pakietu `scikit-learn` [33] w języku Python z następującymi parametrami konstruktora:

- `C=100` - Parametr ten kontroluje siłę regularyzacji, gdzie wartość 100 wskazuje na stosunkowo słabą regularyzację, zwiększającą zdolność modelu do dopasowania się do danych [36].
- `class_weight='balanced'` - Ten parametr zapewnia zrównoważenie wag klas [37].
- `max_iter=10000` - Określa maksymalną liczbę iteracji procesu optymalizacji, co jest niezbędne do osiągnięcia konwergencji w skomplikowanych modelach [38].

- `solver='lbfgs'` - Solver 'lbfgs' jest efektywnym algorytmem kwazi-newtonowskim dla dużych zbiorów danych, który przybliża drugą pochodną funkcji straty, co znacząco przyspiesza proces znajdowania optymalnych parametrów [34].
- `random_state=42` - Ustawienie ziarna losowości zapewnia reprodukowalność wyników.

### Równanie modelu

Model regresji logistycznej dla analizy wieloklasowej można przedstawić jako:

$$p_{ic} = P(y_i = c|X) = \frac{\exp(\beta_{c0} + \beta_{c1}x_{i1} + \dots + \beta_{ck}x_{ik})}{1 + \sum_{c=1}^{C-1} \exp(\beta_{c0} + \beta_{c1}x_{i1} + \dots + \beta_{ck}x_{ik})} \quad (4)$$

gdzie:

- $p_{ic}$  jest prawdopodobieństwem, że obserwacja  $i$  przynależy do klasy  $c$ .
- $y_i$  jest zmienną zależną dla  $i$ -tej obserwacji.
- $X$  reprezentuje macierz zmiennych niezależnych.
- $\beta_{c0}, \beta_{c1}, \dots, \beta_{ck}$  to parametry modelu dla klasy  $c$ , które są estymowane w procesie uczenia, odpowiadające wpływowi poszczególnych zmiennych na prawdopodobieństwo przynależności do danej klasy.
- $x_{i1}, \dots, x_{ik}$  to wartości zmiennych niezależnych dla  $i$ -tej obserwacji, odpowiadające zmiennym, które są brane pod uwagę przy przewidywaniu wyniku.
- Parametr  $C$  w konstruktorze `LogisticRegression` kontroluje siłę regularyzacji, gdzie mniejsza wartość  $C$  oznacza mocniejszą regularyzację, wpływając na wartości estymowanych parametrów  $\beta$ . Mniejsza wartość  $C$  skutkuje mniejszymi wartościami bezwzględnyymi parametrów  $\beta$ , co pomaga zapobiegać nadmiernemu dopasowaniu modelu do danych treningowych.

### 3.2.6. Model Drzew Decyzyjnych

Drzewa decyzyjne (ang. Decision Trees) to narzędzie, które pozwala na modelowanie zarówno zmiennych ciągłych, jak i kategorycznych w problemach klasyfikacji oraz regresji. Są szeroko stosowane ze względu na ich zrozumiałość i zdolność do reprezentowania jasnych decyzji na podstawie atrybutów danych [39].

#### Opis Modelu

Model drzewa decyzyjnego konstruuje hierarchiczną strukturę binarną, w której każdy węzeł wewnętrzny odpowiada punktom decyzyjnym opartym na atrybutach, a liście reprezentują terminalne kategorie wyników klasyfikacji. Proces podziału w węzłach wewnętrznych odbywa się przez selekcję atrybutów, które efektywnie minimalizują wybraną miarę heterogeniczności zbioru danych, tzn. entropię, indeks Giniego, czy błąd klasyfikacji [39], [40]. Decyzje podziału są zwykle podejmowane z wykorzystaniem algorytmów zachłannych, które iteracyjnie dzielą zbiór danych na coraz mniejsze podzbiory, dążąc do maksymalizacji czystości węzłów końcowych, czyli liści. Cały proces prowadzi do systematycznego budowania hierarchicznej

struktury decyzyjnej, która wizualnie przypomina odwrócone drzewo, z korzeniem na górze i liśćmi na dole.

### **Implementacja**

Implementacja w ramach niniejszej pracy korzysta z klasy `DecisionTreeClassifier` z pakietu `scikit-learn` [33] z jednym dodatkowym parametrem:

- `random_state=42`: Gwarantuje stałość wyników poprzez kontrolowany losowy wybór atrybutów.

### **3.2.7. Model Lasów Losowych**

Lasy losowe (ang. Random Forests) to technika, która wykorzystuje koncepcję baggingu (ang. Bootstrap Aggregating) i losowości zmiennych przy wyborze podziałów, by poprawić dokładność i stabilność modeli predykcyjnych. Metoda ta została zaproponowana przez Tin Kam Ho, a następnie ulepszona przez Leo Breimana i opiera się na idei "mądrości tłumów", gdzie agregacja decyzji podjętych przez poszczególne drzewa prowadzi do ulepszenia generalizacji modelu [41], [42].

### **Implementacja**

Model w badaniu został zaimplementowany przy użyciu klasy `RandomForestClassifier` z pakietu `scikit-learn` [33] w języku Python, gdzie konfiguracja parametrów klasyfikatora jest następująca:

- `max_depth=30`: Określa maksymalną głębokość każdego drzewa w lesie, co umożliwia algorytmowi uczenie się bardziej złożonych wzorców w danych.
- `max_features='log2'`: Określa liczbę zmiennych do rozważenia przy poszukiwaniu najlepszego podziału, w tym przypadku będącą logarytmem z podstawą 2 z liczby zmiennych.
- `min_samples_split=2`: Minimalna liczba próbek wymagana do podziału węzła wewnętrznego.
- `n_estimators=500`: Liczba drzew w lesie, co wpływa bezpośrednio na moc predykcyjną i stabilność modelu.
- `class_weight='balanced'`: Wagi klas są przypisywane odwrotnie proporcjonalnie do częstości klas w danych wejściowych.
- `n_jobs=-1`: Instrukcja dla algorytmu do użycia wszystkich dostępnych rdzeni procesora dla obliczeń równoległych.
- `random_state=42`: Parametr zapewniający reprodukowalność wyników poprzez ustalenie ziarna generatora losowości.
- `bootstrap=True`: Używa losowego pobierania próbek z podmianą do budowy drzew, co jest standardowym podejściem w lasach losowych.
- `max_samples=1.0`: Ustawienie tego parametru na 1.0 oznacza, że do budowy każdego drzewa wykorzystywana jest cała próbka danych.

Lasosy losowe redukują wariancję poprzez agregację wielu drzew, minimalizując tym samym ryzyko nadmiernego dopasowania, w przeciwieństwie do pojedynczego drzewa decyzyjnego. Dzięki temu podejściu, są one zdolne do efektywnego radzenia sobie z dużymi zbiorami danych i wysoką wymiarowością przestrzeni zmiennych [43].

### **Bagging (ang. Bootstrap Aggregating)**

Bagging to technika redukcji wariancji, polegająca na tworzeniu wielu wersji modelu na różnych podzbiorach danych, a następnie agregowaniu ich wyników (na przykład przez głosowanie większościowe w klasyfikacji lub średnią w regresji). Każde drzewo w losowym lesie jest trenowane na nieco innej próbkce danych, wygenerowanej za pomocą próbkowania z powtórzeniami (ang. bootstrap) z oryginalnego zestawu danych. To oznacza, że każde drzewo uczy się na nieco innej wersji danych, co pomaga w redukcji ryzyka przeuczenia (overfitting) modelu na danych treningowych [44].

### **Wykorzystanie wszystkich rekordów i części zmiennych w każdym drzewie**

Kluczowym aspektem zastosowanego modelu lasów losowych jest zdefiniowanie parametrów *max\_features* oraz *max\_samples*.

Parametr *max\_features* ustawiony na  $\log_2$  oznacza, że w procesie budowania każdego drzewa decyzyjnego brana jest pod uwagę logarytmiczna liczba (logarytm o podstawie 2) wszystkich dostępnych zmiennych. Dzięki temu podejściu, każde drzewo w lesie wykorzystuje różne kombinacje zmiennych, co przyczynia się do zwiększenia różnorodności w modelu i zmniejszenia ryzyka przeuczenia.

Z kolei parametr *max\_samples* ustawiony na 1.0 jest równoznaczny z wykorzystaniem całego dostępnego zestawu danych do trenowania każdego z drzew. Takie podejście gwarantuje, że każde drzewo uczy się na pełnym zbiorze danych, co zwiększa szansę na dokładne uchwycenie zależności występujących w danych.

## **3.2.8. Model Drzew Wzmacnianych**

Metoda wzmacniania sekwencyjnego drzew decyzyjnych, znana również jako Gradient Boosting, to narzędzie służące do budowania precyzyjnych modeli predykcyjnych przez sekwencyjną agregację wielu słabych klasyfikatorów. Technika ta stosuje iteracyjną korektę błędów poprzednich klasyfikatorów, zwiększając skuteczność modelu przy każdym kroku.

### **Funkcja straty i proces optymalizacji**

Kluczowym elementem jest funkcja straty, która jest minimalizowana w każdej iteracji, co prowadzi do poprawy modelu. Funkcja straty jest definiowana jako:

$$L(y, F(x)) = \sum_{i=1}^n L(y_i, F(x_i)), \quad (5)$$

gdzie:

- $L$  jest miarą błędu predykcji, zwaną funkcją straty, która ocenia różnicę między przewi-



dywanymi wartościami  $F(x_i)$  a rzeczywistymi wartościami  $y_i$ ,

- $y_i$  to rzeczywista etykieta klasy dla  $i$ -tej próbki,
- $x_i$  to zmienne  $i$ -tej próbki,
- $F(x_i)$  to przewidywana wartość przez aktualny model dla  $i$ -tej próbki,
- $n$  to liczba próbek w danych.

W każdej iteracji algorytmu nowe drzewo jest dopasowywane tak, aby jak najlepiej korygować błędy modelu z poprzedniej iteracji. Poprawka ta jest realizowana przez przypisanie większych wag próbkom, które zostały źle sklasyfikowane, co zmusza model do skupienia większej uwagi na trudniejszych przypadkach. Wagi te są aktualizowane na podstawie gradientu funkcji straty, co jest realizowane zgodnie z zasadami optymalizacji gradientowej [45], [46].

### **Implementacja**

Model w badaniu został zaimplementowany poprzez funkcję `GradientBoostingClassifier` z pakietu `scikit-learn` [33] z podstawowymi parametrami.

### **3.2.9. Model Maszyny Wektorów Nośnych**

Model maszyny wektorów nośnych (ang. Support Vector Machine - SVM) jest techniką wykorzystywaną w problemach klasyfikacji i regresji. W kontekście klasyfikacji, SVM znajduje hiperpłaszczyznę w przestrzeni wielowymiarowej, która najlepiej oddziela różne klasy danych. Model ten jest skuteczny, gdy dane są liniowo separowalne, lecz może być również stosowany do danych nieseparowalnych liniowo dzięki funkcjom jądrowym [47], [48].

#### **Matematyczne podstawy**

Centralnym elementem SVM jest maksymalizacja marginesu między klasami danych. Margines to odległość między hiperpłaszczyzną decyzyjną a najbliższymi punktami danych każdej z klas, nazywanymi wektorami nośnymi. Problem optymalizacji modelu jest wyrażony jako problem programowania kwadratowego:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad (6)$$

przy ograniczeniach:

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \text{ dla } i = 1, \dots, n \quad (7)$$

gdzie:

- $\mathbf{w}$  to wektor normalny do hiperpłaszczyzny,
- $b$  to wyraz wolny w równaniu hiperpłaszczyzny,
- $y_i$  to etykiety klas (przyjmują wartości -1 lub 1),
- $\mathbf{x}_i$  to wektory zmiennych.

#### **Funkcje jądrowe**

W sytuacjach, gdy dane nie są liniowo separowalne, SVM używa funkcji jądrowych do mapowania danych wejściowych na przestrzeń o wyższej liczbie wymiarów, gdzie klasy danych mogą być separowane liniowo [49], [48]. Domyślną funkcją jądrową w pakiecie `scikit-learn` jest RBF (Radial Basis Function):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$$
$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (8)$$

gdzie:

- $\gamma$  jest parametrem, który kontroluje szerokość jądra,
- $\|\mathbf{x}_i - \mathbf{x}_j\|$  oznacza normę euklidesową różnicy między wektorami zmiennych.

#### **Implementacja**

W pracy zastosowano SVM z pakietu `scikit-learn` [33]. Zastosowano parametr `probability=True` w celu oszacowania prawdopodobieństwa przynależności do każdej z klas, oraz `random_state=42` w celu zapewnienia reprodukowalności wyników.

### **3.2.10. Klasyfikatory głosujące**

Modelowanie oparte na uczeniu zespołowym jest strategią uczenia maszynowego, podczas której poszczególne modele predykcyjne są łączone w celu osiągnięcia lepszej dokładności i stabilności predykcji. Należące do technik modelowania zespołowego, klasyfikatory głosujące pozwalają na agregację wyników wcześniej stworzonych modeli, wykorzystując strategię głosowania twardego lub miękkiego, aby wybrać ostateczną predykcję klasy [47].

#### **Klasyfikator głosujący twardy**

W podejściu twardym, decyzja klasyfikacyjna jest podejmowana na podstawie najczęściej wybieranej etykiety klasowej przez modele składowe. Jest to podejście oparte na wartości dominującej, które nie bierze pod uwagę niepewności predykcji poszczególnych modeli. Poniżej znajduje się skład zespołu klasyfikatorów zaimplementowanych w pracy do stworzenia modelu głosującego typu twarde:

- Model regresji logistycznej
- Model lasów losowych
- Model drzew decyzyjnych
- Model wzmacnianych drzew decyzyjnych

#### **Klasyfikator głosujący miękki**

W podejściu miękkim, finalna klasa jest determinowana na podstawie średnich ważonych prawdopodobieństw przynależności do klas, przewidywanych przez poszczególne modele. To podejście może lepiej radzić sobie z niepewnościami predykcji i przeuczeniem. Poniżej znajduje się skład klasyfikatorów użytych do stworzenia modelu głosującego typu miękkiego:

- Model regresji logistycznej

- Model wzmacnianych drzew decyzyjnych
- Model SVM
- Model lasów losowych

### Implementacja

W pracy zaimplementowano modele głosujące korzystając z pakietu `scikit-learn` [33].

#### 3.2.11. Istotność zmiennych

W celu dokładnej analizy poszczególnych czynników wpływających na predykcje, zastosowano metodę wyróżnienia istotnych zmiennych.

Proces ten polega na analizie zmian w miarach nieczystości, które są redukowane dzięki podziałom na poszczególnych zmiennych. Podczas budowy drzewa decyzyjnego, algorytm wybiera zmienną, która najlepiej dzieli dane, bazując na kryterium zmniejszenia nieczystości, takim jak wskaźnik Giniego dla klasyfikacji, lub błąd średniokwadratowy dla regresji [50].

Redukcja nieczystości dla podziału opartego na zmiennych jest obliczana według wzoru:

$$\Delta I(f, n) = I(n) - (p_L \times I(n_L) + p_R \times I(n_R)) \quad (9)$$

gdzie:

- $I(n)$  - nieczystość węzła przed podziałem,
- $p_L$  - proporcja próbek w lewym węźle po podziale,
- $I(n_L)$  - nieczystość lewego węzła po podziale,
- $p_R$  - proporcja próbek w prawym węźle po podziale,
- $I(n_R)$  - nieczystość prawego węzła po podziale.

Całkowity wkład zmiennej do modelu, oznaczany jako  $I(i)$ , jest sumą wszystkich redukcji nieczystości, które przypisane są dla tej zmiennej we wszystkich węzłach drzewa [50]:

$$I(i) = \sum_{t \in T} \Delta \text{impurity}(i, t) \quad (10)$$

gdzie:

- $T$  - zbiór wszystkich drzew w modelu,
- $\Delta \text{impurity}(i, t)$  - redukcja nieczystości przypisywana zmiennym w drzewie  $t$ .

Dla modeli takich jak regresja logistyczna, ogólna istotność zmiennej  $i$ , również oznaczana jako  $I(i)$ , jest obliczana jako suma bezwzględnych wartości współczynników wszystkich klas:

$$I(i) = \sum_{k=1}^K |\beta_{ik}| \quad (11)$$

gdzie:

- $K$  - liczba klas,
- $\beta_{ik}$  - współczynnik zmiennej  $i$  dla klasy  $k$ .

### 3.2.12. Metryki dokładności

W celu porównania wyników stworzonych modeli obliczono metryki dokładności opierające się o macierz pomyłek.

Metryki dokładności w modelach klasyfikacyjnych umożliwiają ocenę skuteczności modelu w kontekście przewidywania poprawnych klas. W pracy wykorzystano metryki takie jak dokładność (ang. Accuracy), czułość (ang. Recall), precyzja (ang. Precision), miara F1 (ang. F1-score) oraz pole powierzchni pod krzywą ROC (ang. Receiver Operating Characteristic), wraz z ich zastosowaniem w kontekście wieloklasowym [46].

Macierz pomyłek jest niezbędnym narzędziem do oceny wydajności modeli klasyfikacyjnych, szczególnie w kontekście wieloklasowym. Pozwala ona na szczegółową analizę skuteczności modelu w identyfikacji każdej klasy.

#### Struktura macierzy pomyłek

Dla modelu z czterema klasami: A, B, C, D, macierz pomyłek prezentuje się następująco:

	Przewidziane A	Przewidziane B	Przewidziane C	Przewidziane D
Rzeczywiste A	$TP_A$	$FP_{B \rightarrow A}$	$FP_{C \rightarrow A}$	$FP_{D \rightarrow A}$
Rzeczywiste B	$FN_{A \rightarrow B}$	$TP_B$	$FP_{C \rightarrow B}$	$FP_{D \rightarrow B}$
Rzeczywiste C	$FN_{A \rightarrow C}$	$FN_{B \rightarrow C}$	$TP_C$	$FP_{D \rightarrow C}$
Rzeczywiste D	$FN_{A \rightarrow D}$	$FN_{B \rightarrow D}$	$FN_{C \rightarrow D}$	$TP_D$

(12)

gdzie:

- TP (ang. True Positive, Prawdziwie pozytywne) - Liczba przypadków prawidłowo zaklasyfikowanych jako dana klasa.
- TN (ang. True Negative, Prawdziwie negatywne) - Liczba przypadków prawidłowo zaklasyfikowanych jako nie należące do danej klasy.
- FP (ang. False Positive, Fałszywie pozytywne) - Liczba przypadków niewłaściwie zaklasyfikowanych jako dana klasa.
- FN (ang. False Negative, Fałszywie negatywne) - Liczba przypadków niewłaściwie zaklasyfikowanych jako nie należące do danej klasy.

#### Dokładność (ang. Accuracy)

Dokładność określa stosunek liczby poprawnie zaklasyfikowanych przypadków do całkowitej liczby próbek.

$$\text{Dokładność} = \frac{TP_A + TP_B + TP_C + TP_D}{\text{wszystkie przypadki}} \quad (13)$$

### **Czułość (ang. Recall) dla klasy A**

Czułość dla klasy A mierzy zdolność modelu do prawidłowego zidentyfikowania przypadków klasy A.

$$\text{Czułość dla A} = \frac{TP_A}{TP_A + \sum_{X \neq A} FN_{X \rightarrow A}} \quad (14)$$

### **Precyzja (ang. Precision) dla klasy A**

Precyzja dla klasy A określa stosunek prawidłowo przewidzianych przypadków klasy A do wszystkich przypadków przewidzianych jako klasa A.

$$\text{Precyzja dla A} = \frac{TP_A}{TP_A + \sum_{X \neq A} FP_{X \rightarrow A}} \quad (15)$$

### **Miara F1 dla klasy A**

Miara F1 dla klasy A jest średnią harmonijną czułości i precyzji dla klasy A.

$$F1 \text{ dla A} = 2 \cdot \frac{\text{Precyzja dla A} \cdot \text{Czułość dla A}}{\text{Precyzja dla A} + \text{Czułość dla A}} \quad (16)$$

### **Krzywa ROC (ang. Receiver Operating Characteristic)**

Krzywa ROC jest graficznym przedstawieniem zdolności modelu do różnicowania między klasami. Na osi pionowej umieszcza się czułość, a na osi poziomej (1 - specyficzność), co pozwala ocenić, jak dobrze model rozróżnia dwie klasy (pozytywną i negatywną) przy różnych progach decyzyjnych.

### **Pole powierzchni pod krzywą ROC (ang. Area Under the Curve - AUC)**

AUC mierzy zdolność modelu do różnicowania między klasami. Wartość AUC wynosząca 1.0 wskazuje na doskonałą zdolność klasyfikacji, natomiast wartość 0.5 na brak zdolności dyskryminacyjnej modelu (losowa klasyfikacja). Wartości poniżej 0.5 uznaje się za gorsze od przewidywania losowego.

## **3.2.13. TOPSIS**

W celu ostatecznego porównania wyników i wybrania najlepszego modelu zastosowano technikę TOPSIS (ang. Technique for Order of Preference by Similarity to Ideal Solution) [51].

Metoda TOPSIS jest używana do identyfikacji najbardziej preferowanego rozwiązania przez porównanie każdej z opcji z rozwiązaniem idealnym oraz negatywnie idealnym. Poniżej znajduje się szczegółowy opis kroków tej metody:

### **1. Normalizacja danych**

Dane są normalizowane przy użyciu normy euklidesowej:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}} \quad (17)$$

gdzie:

- $r_{ij}$  - znormalizowana wartość elementu w i-tym wierszu i j-tej kolumnie macierzy decyzyjnej,
- $x_{ij}$  - pierwotna wartość elementu w i-tym wierszu i j-tej kolumnie macierzy decyzyjnej,
- $n$  - liczba alternatyw (wierszy) w macierzy decyzyjnej.

Sumowanie odbywa się po wszystkich wierszach macierzy dla danej kolumny  $j$ , co zapewnia skalowanie każdej kolumny.

## 2. Przypisanie wag dla kryteriów

Wagi  $w_j$  są przypisane dla kryteriów zgodnie z ich znaczeniem:

$$v_{ij} = w_j \times r_{ij} \quad (18)$$

gdzie:

- $v_{ij}$  - wartość elementu w macierzy wagowo znormalizowanej,
- $w_j$  - waga przypisana j-temu kryterium, odzwierciedlająca jego znaczenie,
- $r_{ij}$  - znormalizowana wartość elementu z poprzedniego kroku.

## 3. Określenie rozwiązania idealnego i negatywnie idealnego

Rozwiązania te są definiowane jako:

$$A^+ = \max_i (v_{ij} | j \in J), \quad A^- = \min_i (v_{ij} | j \in J) \quad (19)$$

gdzie:

- $A^+$  - wartość idealna, najlepsza możliwa wartość dla każdego kryterium,
- $A^-$  - wartość negatywnie idealna, najgorsza możliwa wartość dla każdego kryterium,
- $v_{ij}$  - elementy macierzy wagowo znormalizowanej,
- $J$  - zbiór wszystkich kryteriów (kolumn) w macierzy.

## 4. Obliczenie odległości od rozwiązań idealnych

Euklidesowa odległość od rozwiązania idealnego oraz negatywnie idealnego jest obliczana jako:

$$d_i^+ = \sqrt{\sum_{j=1}^m (v_{ij} - A_j^+)^2}, \quad d_i^- = \sqrt{\sum_{j=1}^m (v_{ij} - A_j^-)^2} \quad (20)$$

gdzie:

- $d_i^+$  - odległość euklidesowa od rozwiązania idealnego dla i-tej alternatywy,

- $d_i^-$  - odległość euklidesowa od rozwiązania negatywnie idealnego dla i-tej alternatywy,
- $m$  - liczba kryteriów (kolumn) w macierzy decyzyjnej,
- $v_{ij}$  - elementy macierzy wagowo znormalizowanej,
- $A_j^+, A_j^-$  - wartości idealne i negatywnie idealne dla każdego kryterium.

## 5. Obliczenie wyniku TOPSIS

Wynik dla każdej alternatywy jest obliczany jako:

$$S_i = \frac{d_i^-}{d_i^+ + d_i^-} \quad (21)$$

gdzie:

- $S_i$  - wynik TOPSIS dla i-tej alternatywy,
- $d_i^+, d_i^-$  - odległości od rozwiązania idealnego i negatywnie idealnego.

Alternatywy są rangowane na podstawie wyniku  $S_i$ , gdzie wyższe wartości wskazują na lepsze rozwiązanie.

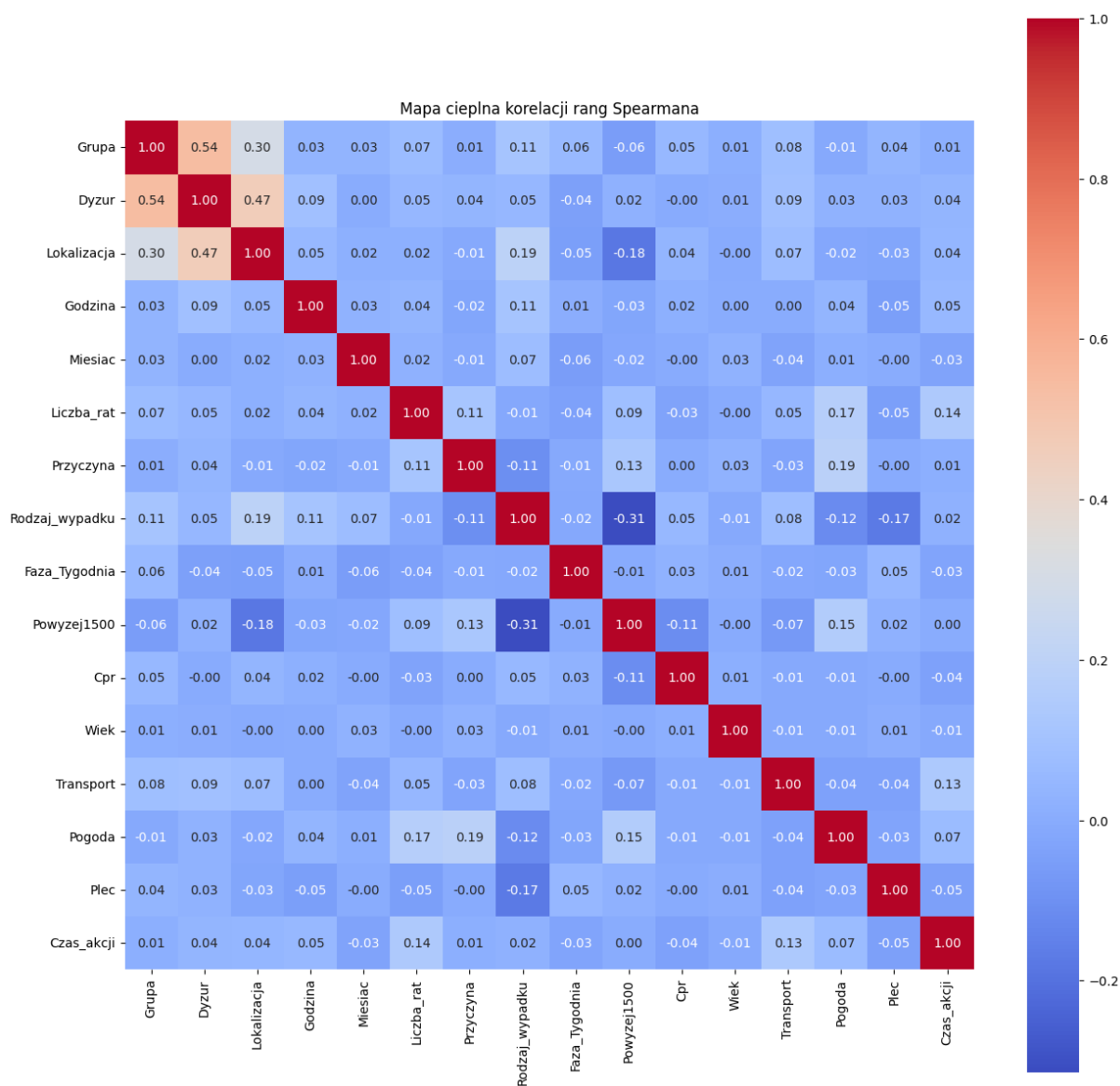
W metodzie TOPSIS istotnym elementem jest przypisanie wag poszczególnym kryteriom, które są brane pod uwagę w procesie podejmowania decyzji. Waga każdego kryterium odzwierciedla jego znaczenie w kontekście całkowitej analizy. Kryteriami metody TOPSIS w przeprowadzonym badaniu były poszczególne metryki dokładności. Z uwagi na istotność prawidłowego przewidywania wypadków śmiertelnych zdecydowano, aby większą wagę przypisać tym metrykom, które najbardziej przyczyniają się do dokładności i wiarygodności przewidywania klasy 3 (zgon). Poniżej znajduje się Tabela z przypisaniem wag poszczególnych metrykom:

Tabela 15: Przypisanie wag metrykom w analizie TOPSIS

Metryka	Waga
Dokładność	1.5
Obszar Pod Krzywą ROC (AUC)	1.5
Czułość Klasy 0	1.5
Czułość Klasy 1	1.0
Czułość Klasy 2	1.0
Czułość Klasy 3	2.0
Miara F1 Klasy 0	1.5
Miara F1 Klasy 1	1.5
Miara F1 Klasy 2	1.5
Miara F1 Klasy 3	2.0
Precyzja Klasy 0	1.5
Precyzja Klasy 1	1.5
Precyzja Klasy 2	1.5
Precyzja Klasy 3	1.5

## 4. Wyniki

W celu doboru odpowiednich zmiennych do modeli, postanowiono obliczyć korelację rang Spearmana, aby wyeliminować zmienne wysoko skorelowane. Na Rysunku 1 przedstawiono macierz korelacji w formie mapy cieplnej.

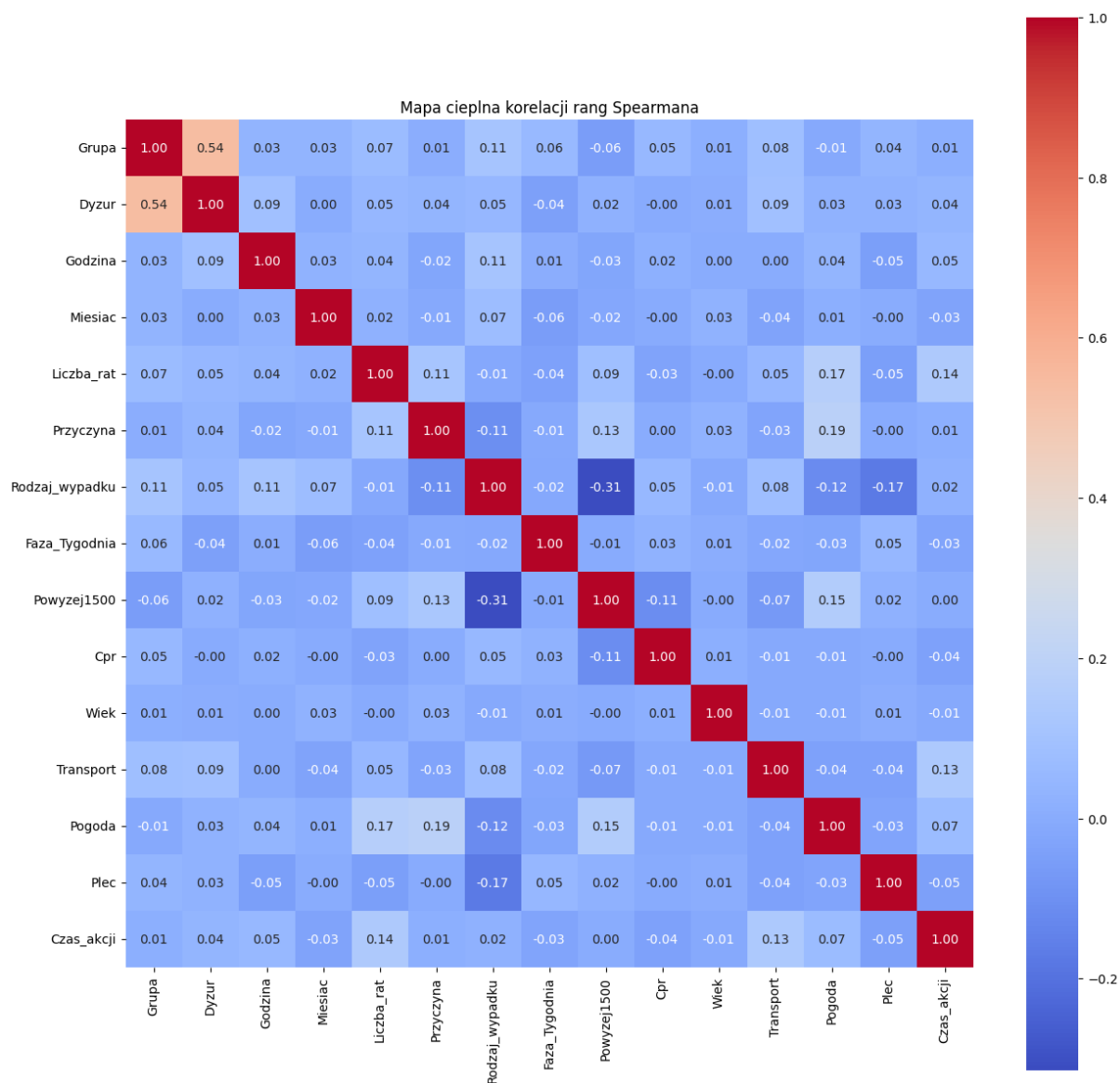


Rysunek 1: Mapa cieplna korelacji rang Spearmana

Na przekątnej, w intensywnie czerwonym kolorze, przedstawione są wartości korelacji dla identycznych zmiennych. Wartość 1,00 oraz głęboki odcień czerwieni symbolizują silną korelację pozytywną, co oznacza, że wzrost wartości jednej zmiennej pociąga za sobą wzrost wartości drugiej. Kolorem niebieskim zaznaczono korelację negatywną, wskazującą na to, że wzrost wartości jednej zmiennej skutkuje spadkiem wartości drugiej. Biały kolor wskazuje na brak korelacji lub korelację bardzo słabą.

Najwyższe wartości korelacji zaobserwowano między zmiennymi *Dyzur* i *Grupa* (0,54),





Rysunek 2: Mapa cieplna korelacji rang Spearmana

a także między *Dyżur* i *Lokalizacja* (0,47), co sugeruje umiarkowaną korelację pozytywną. Na granicy między korelacją umiarkowaną a słabą znajduje się związek między *Lokalizacja* a *Grupa* (0,30) oraz między *Rodzaj wypadku* a *Powyżej 1500 m.n.p.m.* (-0,31), gdzie zaobserwowano korelację ujemną.

W procesie optymalizacji modelu analizowano wpływ poszczególnych zmiennych na wyniki. Zmienne takie jak *Dyżur*, *Grupa* i *Lokalizacja* były sukcesywnie usuwane z modelu, w celu oceny ich wpływu. Najlepsze rezultaty osiągnięto po usunięciu zmiennej *Lokalizacja*. Pomimo umiarkowanej korelacji między zmiennymi *Dyżur* i *Grupa*, zdecydowano o ich zachowaniu w analizie, gdyż eliminacja którejkolwiek z nich negatywnie wpływała na wyniki. Na Rysunku 2 przedstawiono finalną mapę cieplną korelacji rang Spearmana.

W celu optymalizacji wydajności modeli, podjęto decyzję o zastosowaniu kodowania jednoelementowanego (ang. one-hot encoding) dla danych katégorycznych. Procedura ta polega na

przekształceniu zmiennych kategoriycznych na wartości binarne 0 i 1, co prowadzi do zwiększenia ogólnej liczby zmiennych w analizowanym zbiorze danych. Finalnie uzyskano 110 zmiennych.

Cały zbiór danych został podzielony na zestaw treningowy i testowy w proporcjach 0.75 (zbiór treningowy) oraz 0.25 (zbiór testowy).

Następnym krokiem było wybranie optymalnego zestawu zmiennych, które pozwolą na uzyskanie najlepszych wyników analizy. Użyto do tego techniki RFECV z domyślnym klasyfikatorem *GradientBoostingClassifier* oraz 3 warstwami walidacji krzyżowej. Do analizy użyto wszystkich 110 zmiennych. Na koniec procesu RFECV każdej zmiennej przypisywana jest wartość rankingu, gdzie zmienne z niższą wartością rankingu są uznane za bardziej wartościowe. Ranking 1 oznacza, że zmienna była obecna w końcowym modelu, natomiast wyższe wartości wskazują na wcześniejsze usunięcie zmiennej w procesie eliminacji. Wyniki przedstawiono w Tabeli 16.

Tabela 16: Wyniki RFECV

ID	Zmienna i ranking	ID	Zmienna i ranking
1	Godzina: 1	56	Dyzur_Sokolec: 1
2	Liczba_rat: 1	57	Dyzur_Stacja Centralna: 1
3	Wiek: 1	58	Dyzur_Stacja Centralna - Dzień: 1
4	Czas_akcji: 1	59	Dyzur_Stacja Centralna Sanok - DZIEŃ: 19
5	Grupa_Beskidzka: 1	60	Dyzur_Stacja Centralna Sanok - NOC: 46
6	Grupa_Bieszczadzka: 1	61	Dyzur_Stacja Karpacz: 1
7	Grupa_Jurajska: 1	62	Dyzur_Stacja Szklarska Poręba: 1
8	Grupa_Karkonoska: 14	63	Dyzur_Stacja Świeradów Zdrój: 32
9	Grupa_Krynica: 1	64	Dyzur_Szczawnica: 1
10	Grupa_Podhalańska: 1	65	Dyzur_TPR Banica: 51
11	Grupa_Sudecka: 38	66	Dyzur_TSR Jaworzyna: 30
12	Dyzur_Beskid Mały: 45	67	Dyzur_TSR Krempna - MPN: 44
13	Dyzur_Beskid Śląski: 1	68	Dyzur_TSR Prehyba: 47
14	Dyzur_CSR Krynica - DZIEŃ: 31	69	Dyzur_Turbacz Dyżurka: 1
15	Dyzur_CSR Krynica-NOC: 20	70	Dyzur_WISŁA: 34
16	Dyzur_CSR Rabka Zdrój - DZIEŃ: 4	71	Dyzur_Zieleniec: 41
17	Dyzur_CSR Rabka Zdrój - DZIEŃ, Szczawnica: 5	72	Dyzur_zabezpieczenie zawodów sportowych: 33
18	Dyzur_CSR Rabka Zdrój - DZIEŃ, Turbacz Dyżurka: 48	73	Dyzur_Śnieżnica schronisko: 11
19	Dyzur_CSR Rabka Zdrój - NOC: 40	74	Dyzur_ŻAR - LATO: 21
20	Dyzur_CSR Szczyrk - DZIEŃ: 1	75	Miesiac_1: 1
21	Dyzur_CSR Szczyrk - DZIEŃ, Hala Skrzyczneńska: 49	76	Miesiac_2: 1
22	Dyzur_CSR Szczyrk - NOC: 43	77	Miesiac_3: 3
23	Dyzur_Centrala - Jelenia Góra - DZIEŃ- Karpacz: 1	78	Miesiac_4: 1
24	Dyzur_Centrala - Jelenia Góra-DZIEŃ-Karpacz, Stacja Szklarska Poręba: 22	79	Miesiac_5: 18
25	Dyzur_Centrala - Jelenia Góra - NOC / KARPACZ: 36	80	Miesiac_6: 1
26	Dyzur_Centrala - Jelenia Góra - NOC / KARPACZ, Dyspozycja: 53	81	Miesiac_7: 8
27	Dyzur_DS Dukla - DZIEŃ: 26	82	Miesiac_8: 1
28	Dyzur_Dolinki - Dzień: 17	83	Miesiac_9: 23
29	Dyzur_Dyspozycja: 15	84	Miesiac_10: 7
30	Dyzur_Karłów: 50	85	Miesiac_11: 1
31	Dyzur_Krościenko: 24	86	Miesiac_12: 37
32	Dyzur_Krościenko, Szczawnica: 42	87	Przyczyna_Brak_objektywnej: 1
33	Dyzur_Masyw Babiej Góry: 1	88	Przyczyna_Brak_wiedzy/sprzetu: 1
34	Dyzur_Międzygórze: 9	89	Przyczyna_Inne: 1
35	Dyzur_Nowy Targ - Długa Polana: 1	90	Przyczyna_Lawina: 28
36	Dyzur_Nowy Targ - Długa Polana, Turbacz Dyżurka: 2	91	Przyczyna_Niska_temp: 1
37	Dyzur_Ochocka: 25	92	Przyczyna_Upadek: 1
38	Dyzur_Olsztyn - Dzień: 35	93	Przyczyna_Upadek_wysokosc: 1
39	Dyzur_Pasmo Piłska: 1	94	Przyczyna_Wysoka_temp: 10
40	Dyzur_Przysłop: 27	95	Przyczyna_Zabłaznienie: 1
41	Dyzur_RSR Cisna - DZIEŃ: 13	96	Przyczyna_Zachorowanie: 1
42	Dyzur_RSR Piwniczna: 6	97	Przyczyna_ukoszenie/pogryzienie: 1
43	Dyzur_RSR Ustrzyki Górze - DZIEŃ: 1	98	Rodzaj_wypadku_Ekstremalne: 1
44	Dyzur_RSR Ustrzyki Górze - NOC: 39	99	Rodzaj_wypadku_Inne: 1
45	Dyzur_Sokolec: 1	100	Rodzaj_wypadku_Narciarski: 1
46	Dyzur_Stacja Centralna: 1	101	Rodzaj_wypadku_Prace_Lesne: 1
47	Dyzur_Stacja Centralna - Dzień: 1	102	Rodzaj_wypadku_Rowerowy: 1
48	Dyzur_Stacja Centralna Sanok - DZIEŃ: 19	103	Rodzaj_wypadku_Samobójstwo: 1
49	Dyzur_Stacja Centralna Sanok - NOC: 46	104	Rodzaj_wypadku_Trening_sportowy: 1

Kontynuacja na następnej stronie

Tabela 16 – kontynuacja z poprzedniej strony

ID	Zmienna i ranking	ID	Zmienna i ranking
50	Dyzur_Stacja Karpacz: 1	105	Rodzaj_wypadku_Turystyczny: 1
51	Dyzur_Stacja Szklarska Poręba: 1	106	Rodzaj_wypadku_Zbedny_alarm: 1
52	Dyzur_Stacja Świeradów Zdrój: 32	107	Faza_Tygodnia_1: 1
53	Dyzur_Szczawnica: 1	108	Faza_Tygodnia_2: 12
54	Dyzur_TPR Banica: 51	109	Powyzej1500_0: 1
55	Dyzur_TSR Jaworzyna: 30	110	Powyzej1500_1: 1

Zastosowanie techniki RFECV pozwoliło na wyselekcjonowanie 58 najbardziej istotnych zmiennych, które zostały wykorzystane do dalszej analizy. Wybrane zmienne zostały przedstawione w Tabeli poniżej:

Tabela 17: Wyselekcjonowane zmienne na podstawie RFECV

ID	Nazwa zmiennej	ID	Nazwa zmiennej
1	Godzina	31	Przyczyna_Brak_wiedzy/sprzetu
2	Liczba_rat	32	Przyczyna_Inne
3	Wiek	33	Przyczyna_Niska_temp
4	Czas_akcji	34	Przyczyna_Upadek
5	Grupa_Beskidzka	35	Przyczyna_Upadek_wysokosc
6	Grupa_Bieszczadzka	36	Przyczyna_Zabladzenie
7	Grupa_Jurajska	37	Przyczyna_Zachorowanie
8	Grupa_Krynicka	38	Przyczyna_ukoszenie/pogryzienie
9	Grupa_Podhalańska	39	Rodzaj_wypadku_Ekstremalne
10	Dyzur_Beskid Śląski	40	Rodzaj_wypadku_Inne
11	Dyzur_CSR Szczyrk - DZIEN	41	Rodzaj_wypadku_Narciarski
12	Dyzur_Centrala - Jelenia Góra - DZIEN - Karpacz	42	Rodzaj_wypadku_Prace_Lesne
13	Dyzur_Masyw Babiej Góry	43	Rodzaj_wypadku_Rowerowy
14	Dyzur_Nowý Targ - Długa Polana	44	Rodzaj_wypadku_Samobójstwo
15	Dyzur_Pasmo Piłska	45	Rodzaj_wypadku_Trening_sportowy
16	Dyzur_RSR Ustrzyki Górne - DZIEN	46	Rodzaj_wypadku_Turystyczny
17	Dyzur_Sokolec	47	Rodzaj_wypadku_Zbedny_alarm
18	Dyzur_Stacja Centralna	48	Faza_Tygodnia_1
19	Dyzur_Stacja Centralna - Dzień	49	Powyzej1500_0
20	Dyzur_Stacja Karpacz	50	Powyzej1500_1
21	Dyzur_Stacja Szklarska Poręba	51	Cpr_0
22	Dyzur_Szczawnica	52	Cpr_1
23	Dyzur_Turbacz Dyżurka	53	Transport_Nie_dotyczny
24	Miesiac_1	54	Transport_Pojazd
25	Miesiac_2	55	Transport_Ręczny
26	Miesiac_4	56	Transport_Smigłowiec
27	Miesiac_6	57	Pogoda_Dobra
28	Miesiac_8	58	Plec_m
29	Miesiac_11		
30	Przyczyna_Brak_obiektywnej		

Kolejnym krokiem było użycie metody SMOTE do zbalansowania klas zmiennej przewidywanej. Poniżej przedstawiono Tabelę z rozkładem zmiennej NACA (dla zbioru treningowego) po podziale danych na zbiór treningowy i testowy.

Tabela 18: Rozkład zmiennej NACA w zbiorze treningowym

NACA	Liczba przypadków
0	1996
Kontynuacja na następnej stronie	

**Tabela 18 – kontynuacja z poprzedniej strony**

NACA	Liczba przypadków
1	2329
2	2630
3	138

Zastosowanie techniki SMOTE pozwoliło na zwiększenie liczebności każdej z klas do 2630 przypadków.

Wszystkie klasyfikatory zostały przetrenowane i przetestowane na tym samym zbiorze treningowym i testowym. Obliczone metryki dokładności znajdują się w Tabelach 19 oraz 20. Brak wyniku AUC w Tabeli 19 dla klasyfikatora głosującego twardego wynika ze sposobu jego działania. Model ten nie dostarcza prawdopodobieństw przynależności do poszczególnych klas, uniemożliwiając obliczenie krzywej ROC i tym samym AUC. W celu poprawy czytelności tabel, w opisie modeli zastosowano ich skróty wyszczególnione poniżej:

- **RL** - Model regresji logistycznej
- **LL** - Model lasów losowych
- **DD** - Model drzew decyzyjnych
- **SVM** - Model maszyny wektorów nośnych
- **DW** - Model drzew wzmacnianych
- **KGT** - Model klasyfikatora głosującego twardego
- **KGM** - Model klasyfikatora głosującego miękkiego

**Tabela 19: Porównanie wyników dla wszystkich klas: Dokładność, AUC, Czułość**

Model	Dokładność	AUC	Czułość kl. 0	Czułość kl. 1	Czułość kl. 2	Czułość kl. 3
RL	0.66	0.86	0.77	0.51	0.71	0.43
LL	0.69	0.88	0.84	0.54	0.72	0.24
DD	0.58	0.67	0.70	0.49	0.59	0.17
SVM	0.42	0.71	0.17	0.42	0.60	0.74
DW	0.67	0.87	0.80	0.53	0.71	0.46
KGT	0.68	-	0.86	0.54	0.67	0.30
KGM	0.68	0.88	0.81	0.54	0.71	0.43

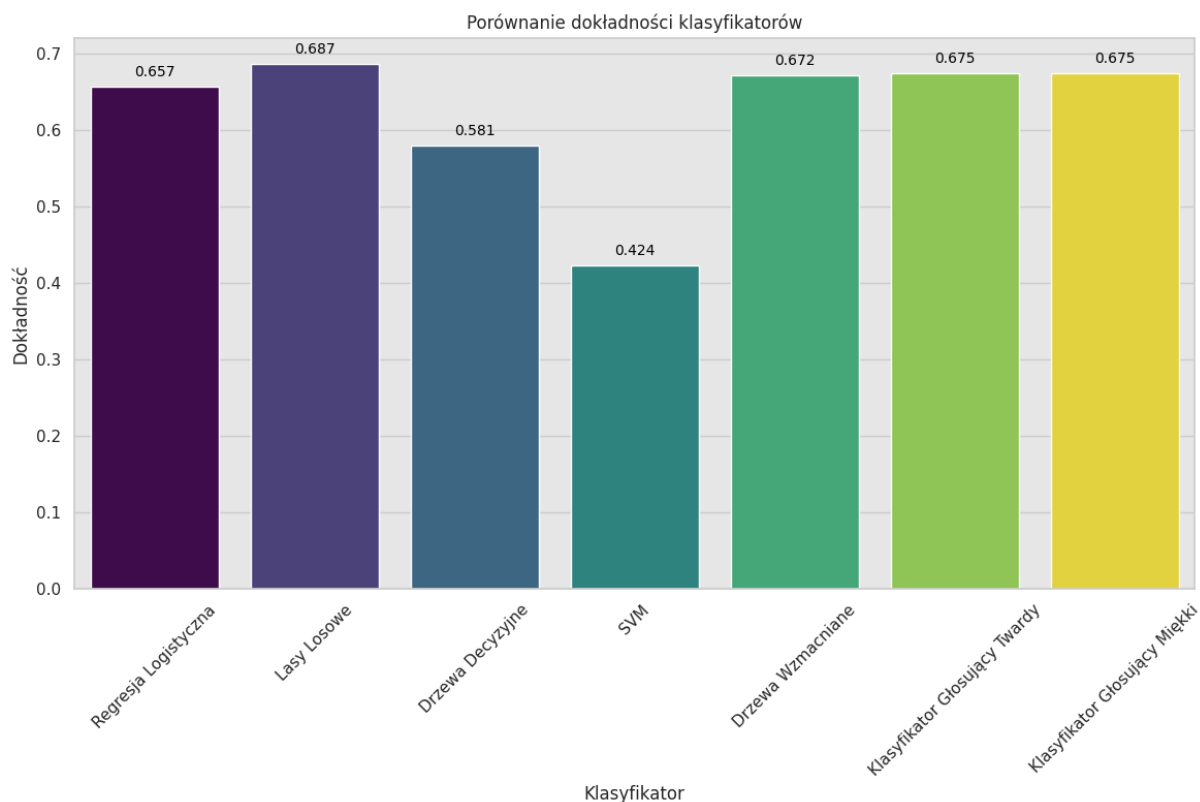
**Tabela 20: Porównanie wyników dla wszystkich klas: Precyzja i Miara F1**

Model	Precyzja kl. 0	F1 kl. 0	Precyzja kl. 1	F1 kl. 1	Precyzja kl. 2	F1 kl. 2	Precyzja kl. 3	F1 kl. 3
RL	0.72	0.74	0.60	0.55	0.68	0.70	0.28	0.34
LL	0.74	0.79	0.62	0.58	0.71	0.71	0.37	0.29
DD	0.66	0.68	0.48	0.49	0.63	0.61	0.16	0.16
SVM	0.52	0.25	0.52	0.47	0.46	0.52	0.09	0.17
DW	0.72	0.76	0.60	0.57	0.71	0.71	0.33	0.38
KGT	0.70	0.77	0.60	0.57	0.72	0.70	0.39	0.34
KGM	0.73	0.77	0.61	0.57	0.71	0.71	0.28	0.34

Wyniki modeli wykazały, że większość zastosowanych klasyfikatorów osiągnęła najlepsze wyniki predykcyjne dla klasy 0 (brak obrażeń), a następnie dla klasy 2 (obrażenia średnie i ciężkie), klasy 1 (obrażenia lekkie) i najmniej skutecznie dla klasy 3 (zgon). Klasyfikator SVM

był wyjątkiem, ponieważ najlepiej radził sobie z predykcją klas 2 i 1, jednak gorzej poradził sobie z klasami 0 i 3. Mimo to, klasyfikator ten zidentyfikował największą liczbę rzeczywistych zgonów w zbiorze testowym (74%), choć kosztem niskiej precyzji (17%).

Na Rysunku 3 przedstawiono porównanie dokładności klasyfikatorów.



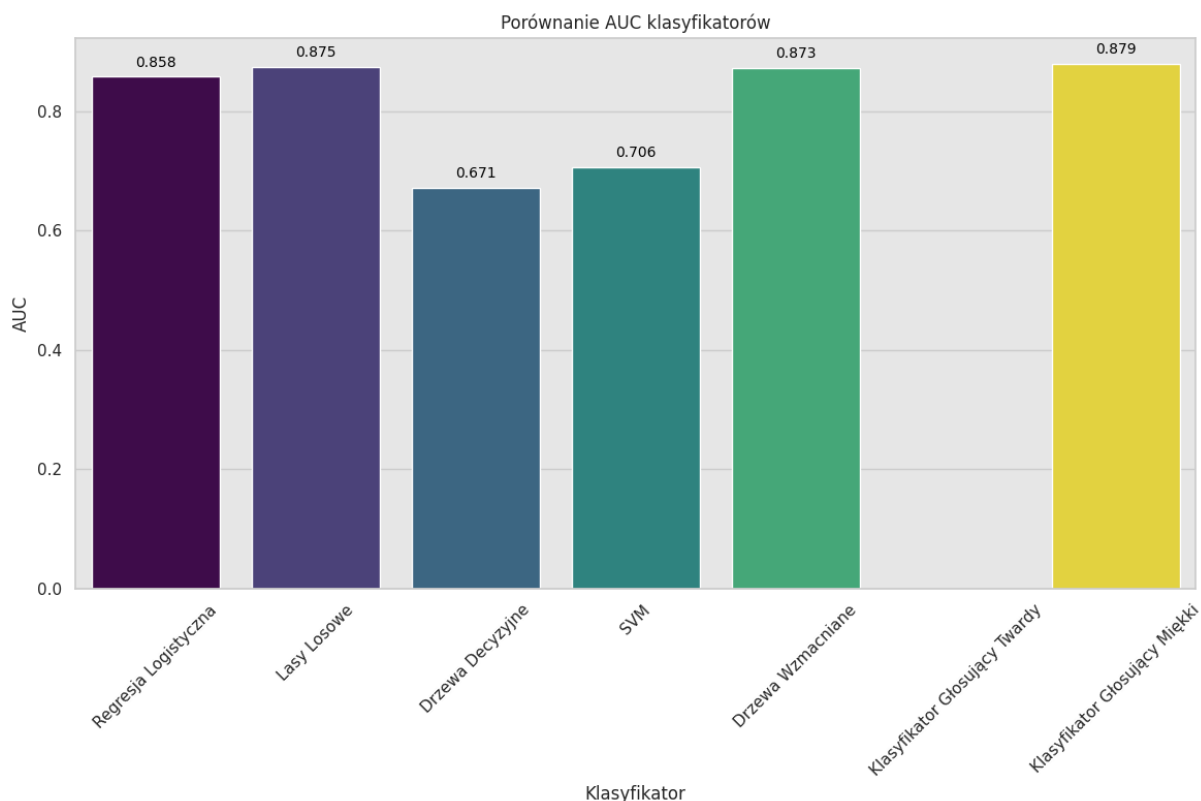
Rysunek 3: Porównanie dokładności klasyfikatorów

Z wykresu wynika, że najwyższą dokładność osiągnął model *Lasów Losowych*, uzyskując wynik **0.687**, co oznacza, że poprawnie przewidział 68.7% wszystkich odpowiedzi. Kolejne były modele *Klasyfikatora Głosującego Miękiego* oraz *Klasyfikatora Głosującego Twardego*, oba osiągające dokładność **0.675**. Model *Drzew Wzmacnianych* osiągnął podobny wynik **0.672**. Najniższą dokładność zanotował model *SVM*, uzyskując wartość **0.424**.

Biorąc pod uwagę ogólną dokładność przewidywań, model *Lasów Losowych* radzi sobie najlepiej, osiągając najwyższą wartość. Niemniej jednak, różnica między nim a trzema kolejnymi modelami – *Klasyfikatorami Głosującymi (Miękkim i Twardym)* oraz *Drzewami Wzmacnianymi* – jest niewielka. Wartości te wskazują, że wszystkie wymienione modele wykazują podobną skuteczność w przewidywaniu wyników.

Na Rysunku 4 przedstawiono porównanie wyników AUC dla stworzonych modeli.

Z wykresu wynika, że najwyższą wartość AUC osiągnęły modele *Klasyfikatora Głosującego Miękiego*, *Lasów Losowych* oraz *Drzew Wzmacnianych*, uzyskując wartości odpowiednio **0.879**, **0.875** i **0.873**. Oznacza to, że te modele mają najwyższą zdolność do prawidłowego rozróżniania między pozytywnymi i negatywnymi przypadkami. Model *Regresji Logistycznej*



Rysunek 4: Porównanie AUC klasyfikatorów

również osiągnął wysoką wartość AUC, wynoszącą **0.858**. Najniższą wartość AUC zanotował model *Drzew Decyzyjnych*, uzyskując wynik **0.671**.

Podczas oceny klasyfikatorów zwrócono szczególną uwagę na jakość przewidywań klasy 3 (zgon), gdyż zdarzenia śmiertelne są najbardziej niepożądanym wynikiem wypadków.

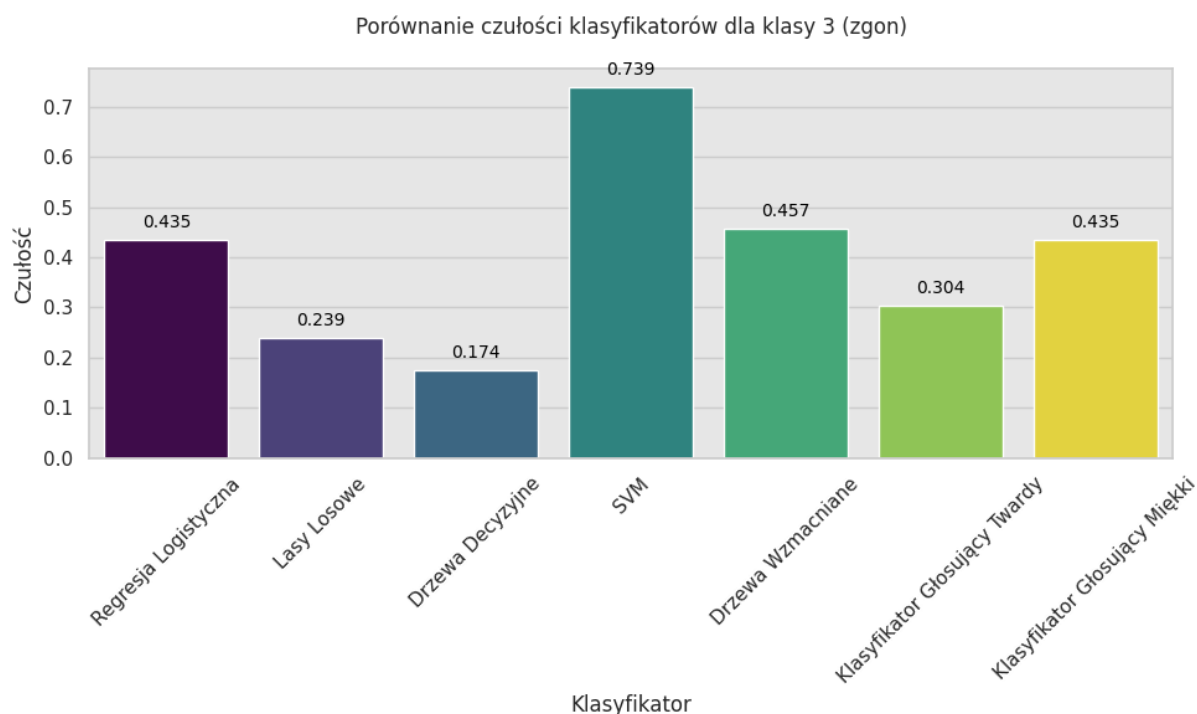
Na Rysunku 5 przedstawiono porównanie czułości klasyfikatorów dla klasy 3 (zgon).

Z wykresu wynika, że najwyższą czułość dla klasy 3 osiągnął model *SVM*, uzyskując wartość **0.739**. Oznacza to, że model wykrył 73.9% prawdziwych wypadków śmiertelnych w zbiorze testowym. *Drzewa Wzmacniane* również wykazały stosunkowo wysoką czułość, osiągając wartość **0.457**. Modele *Regresji Logistycznej* oraz *Klasyfikatora Głosującego Miękkiego* osiągnęły identyczną wartość czułości na poziomie **0.435**. Najniższą czułość zanotował model *Drzew Decyzyjnych*, uzyskując wynik **0.174**.

Pomimo najniższej dokładności (**0.424**) i drugiego najniższego wyniku AUC (**0.706**), model *SVM* wykazuje najwyższą czułość (**0.739**), co oznacza, że najlepiej identyfikuje prawdziwe przypadki zgonów w danych testowych.

Model *Lasów Losowych*, który osiągał najwyższe wyniki dokładności (**0.687**) i AUC (**0.875**), stosunkowo słabo radzi sobie z wykrywaniem prawdziwych wypadków śmiertelnych, osiągając czułość na poziomie **0.239**.

Model *Drzew Wzmacnianych*, osiągając drugą najwyższą wartość czułości (**0.457**), osiąga również wysokie wyniki w pozostałych metrykach (dokładność - **0.672**, AUC - **0.873**), co



Rysunek 5: Porównanie czułości klasyfikatorów dla klasy 3 (zgon)

sugeruje jego stabilność i skuteczność w przewidywaniach.

Na Rysunku 6 przedstawiono porównanie precyzji klasyfikatorów dla klasy 3 (zgon).

Z wykresu wynika, że najwyższą precyzję osiągnął *Klasyfikator Głosujący Twardy*, uzyskując wartość **0.389**, co oznacza, że 38.9% przypadków oznaczonych przez model jako wypadki śmiertelne rzeczywiście nimi było. Kolejnymi modelami o wysokiej precyzji są *Lasy Losowe* (**0.367**) oraz *Drzewa Wzmacniane* (**0.328**).

Model *SVM* osiągnął najniższą precyzję (**0.094**), co wskazuje, że tylko 9.4% przypadków oznaczonych przez niego jako wypadki śmiertelne rzeczywiście nimi było. Model *Drzew Decyzyjnych* również wykazuje niską precyzję (**0.157**).

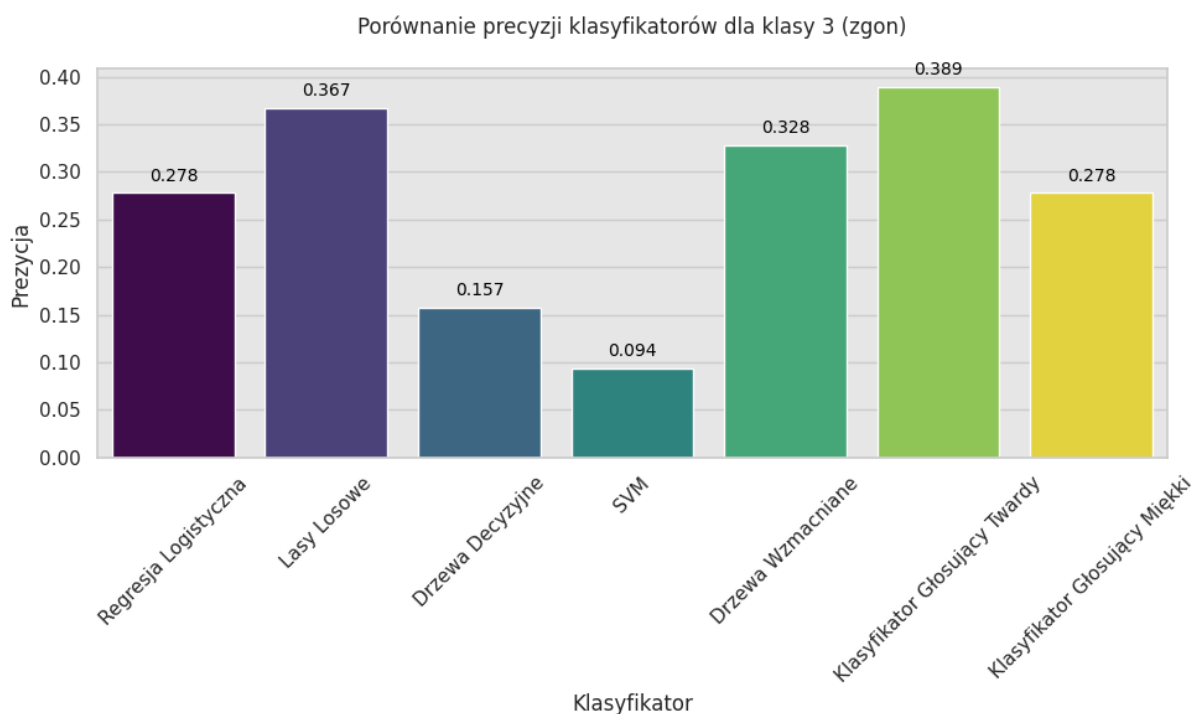
*Klasyfikator Głosujący Twardy* osiąga najwyższą precyzję (**0.389**), co oznacza, że najlepiej identyfikuje prawdziwe zgony spośród wszystkich zgłoszonych jako zgony. Jednak jego czułość (**0.304**) jest relatywnie niska, co sugeruje, że nie wykrywa wielu rzeczywistych przypadków zgonów.

*Lasy Losowe* mają wysoką precyzję (**0.367**), ale niską czułość (**0.239**), co oznacza, że dobrze identyfikują prawdziwe przypadki, ale mogą przegapiać wiele rzeczywistych zgonów.

*Drzewa Wzmacniane* osiągają dobrą równowagę pomiędzy precyzją (**0.328**) a czułością (**0.457**), co czyni je wszechstronnym wyborem.

*SVM* osiąga najwyższą czułość (**0.739**), co oznacza, że najlepiej wykrywa rzeczywiste zgony, ale ma najniższą precyzję (**0.094**), wskazując na dużą liczbę fałszywych pozytywnych.

*Regresja Logistyczna* osiąga umiarkowaną precyzję (**0.278**) oraz umiarkowaną czułość (**0.435**). Oznacza to, że model ten jest stosunkowo zrównoważony, ale nie wyróżnia się w żadnej



Rysunek 6: Porównanie precyzji klasyfikatorów dla klasy 3 (zgon)

z metryk.

Na Rysunku 7 przedstawiono porównanie wyniku miary F1 klasyfikatorów dla klasy 3 (zgon).

*Drzewa Wzmocnione* osiągnęły najwyższą wartość miary F1 (**0.382**), co sugeruje, że model ten oferuje najlepszą równowagę między precyzją a czułością.

*Klasyfikator Głoszący Twardy* (**0.341**), *Klasyfikator Głoszący Miękki* (**0.339**) i *Regresja Logistyczna* (**0.339**) również wykazują wysokie wyniki miary F1, co czyni je efektywnymi w identyfikacji przypadków śmiertelnych.

*SVM* (**0.167**) i *Drzewa Decyzyjne* (**0.165**) osiągnęły najniższe wyniki miary F1, co wskazuje na ich ograniczoną skuteczność w identyfikacji przypadków klasy 3.

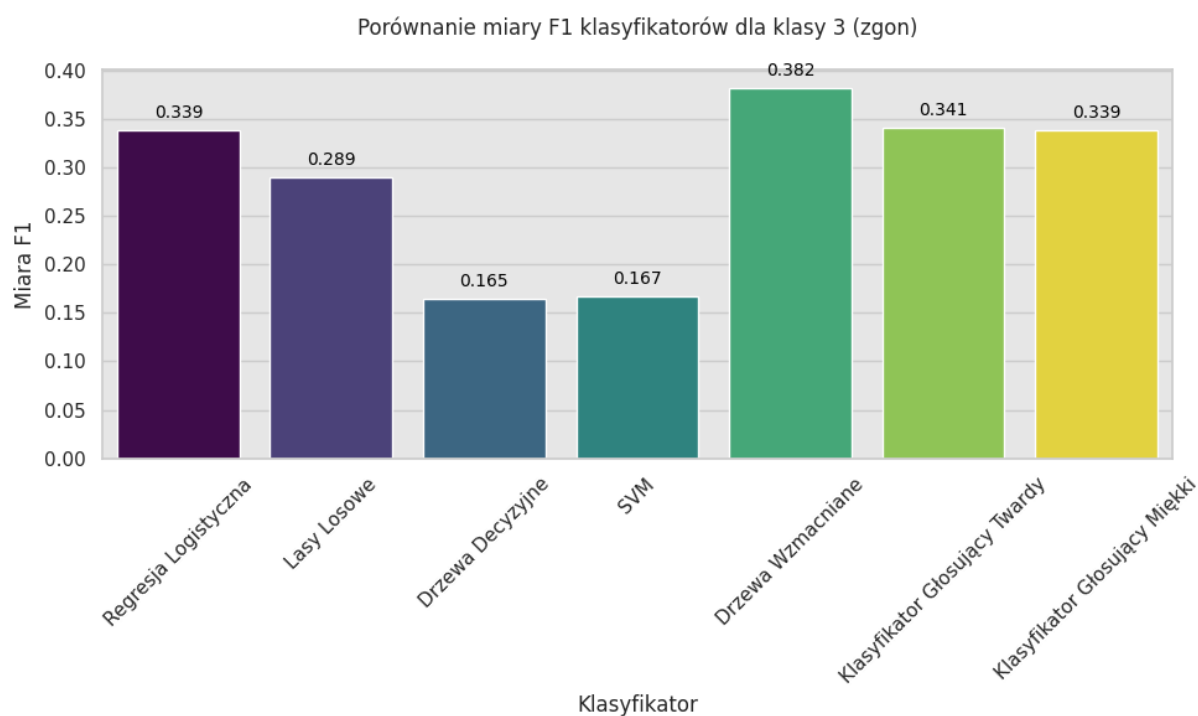
Na Rysunku 8 przedstawiono porównanie skuteczności sześciu różnych klasyfikatorów w rozróżnianiu czterech klas w danych testowych. Każdy wykres ROC ilustruje zdolność modelu do odróżniania prawdziwych pozytywów od fałszywych pozytywów, mierzoną za pomocą współczynnika AUC.

Model *Regresji Logistycznej* wykazuje wysoką skuteczność w rozróżnianiu klasy 0 (AUC = **0.92**) i klasy 3 (AUC = **0.92**). Klasa 2 również osiąga dobry wynik (AUC = **0.85**), natomiast klasa 1 uzyskuje umiarkowaną skuteczność (AUC = **0.77**).

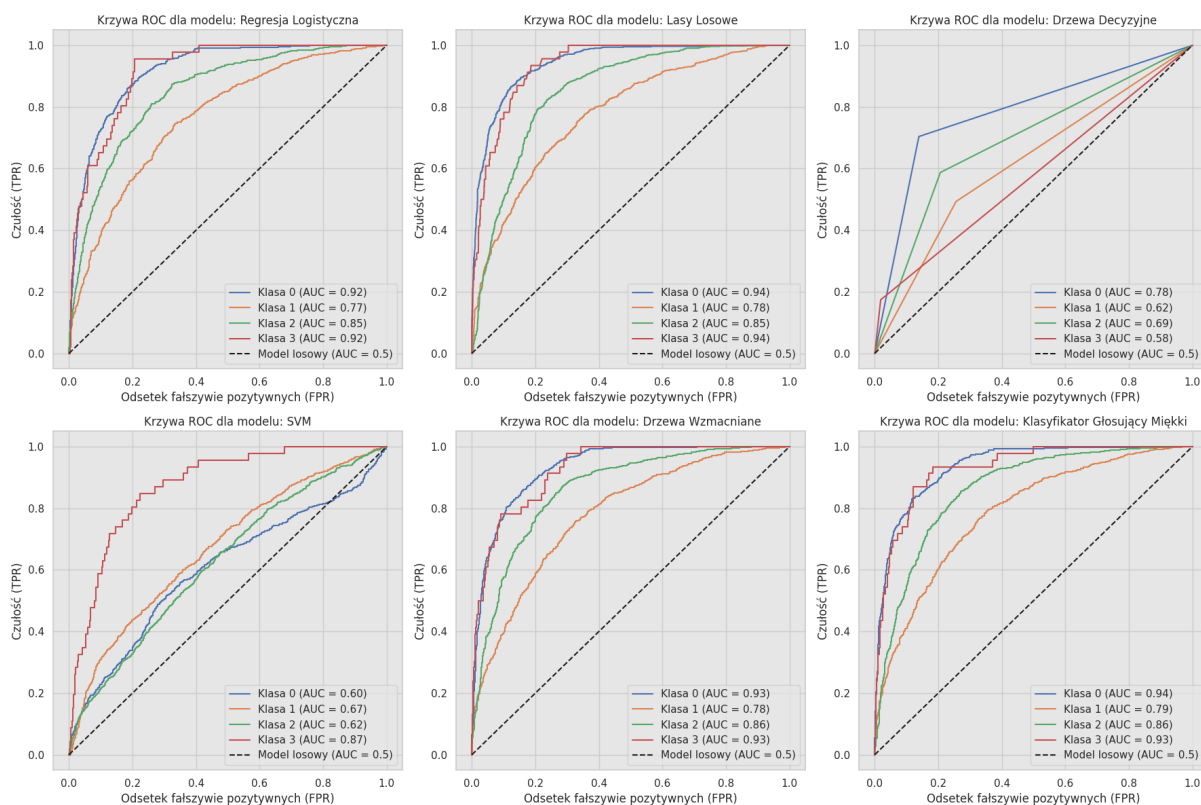
*Lasy Losowe* osiągają najwyższe wartości AUC dla klasy 3 (AUC = **0.94**) oraz klasy 0 (AUC = **0.94**). Klasa 2 ma dobry wynik (AUC = **0.85**), natomiast klasa 1 osiąga umiarkowaną skuteczność (AUC = **0.78**).

Model *Drzew Decyzyjnych* uzyskuje najlepszy wynik AUC dla klasy 0 (AUC = **0.78**).





Rysunek 7: Porównanie wyniku miary F1 klasyfikatorów dla klasy 3 (zgon)



Rysunek 8: Porównanie wykresów krzywych ROC dla każdej klasy

Wyniki dla pozostałych klas są stosunkowo niskie, odpowiednio **0.58**, **0.62** i **0.69** dla klasy 3, klasy 1 i klasy 2.

Model *SVM* osiąga najwyższe AUC dla klasy 3 (AUC = **0.87**). Pozostałe klasy mają niższe wartości AUC, odpowiednio **0.60**, **0.62** i **0.67** dla klasy 0, klasy 2 i klasy 1.

*Drzewa Wzmacniane* wykazują wysoką skuteczność w rozróżnianiu klasy 3 (AUC = **0.93**) i klasy 0 (AUC = **0.93**). Klasa 2 osiąga dobry wynik (AUC = **0.86**), natomiast klasa 1 ma umiarkowaną skuteczność (AUC = **0.78**).

*Klasyfikator Głosujący Miękki* osiąga najwyższe wartości AUC dla klasy 0 (AUC = **0.94**) oraz klasy 3 (AUC = **0.93**). Klasa 2 uzyskuje dobry wynik (AUC = **0.86**), natomiast klasa 1 osiąga umiarkowaną skuteczność (AUC = **0.79**).

*Lasy Losowe*, *Drzewa Wzmacniane*, *Regresja Logistyczna* oraz *Klasyfikator Głosujący Miękki* wykazują najwyższą skuteczność w rozróżnianiu klasy 3 i klasy 0, co czyni je najbardziej efektywnymi modelami w kontekście tych klas. Prezentują także wysokie wartości AUC dla klasy 1 i 2, co łącznie daje wysokie i stabilne wyniki.

*Drzewa Decyzyjne* i *SVM* charakteryzują się niższymi wartościami AUC, co wskazuje na ich ograniczoną skuteczność w klasyfikacji wszystkich klas.

Rysunek 9 przedstawia porównanie krzywych ROC dla trzech wybranych klasyfikatorów: *Drzew Wzmacnianych*, *Klasyfikatora Głosującego Miękkiego* oraz *Regresji Logistycznej*. Ze względu na najlepsze ogólne wyniki metryk precyzji, czułości oraz AUC, do porównania wybrano te trzy klasyfikatory.

Rysunek 9 pokazuje, że wszystkie trzy modele są skuteczne w klasyfikacji, z niewielkimi różnicami w wydajności dla poszczególnych klas.

Na Rysunku 10 przedstawiono wykresy wskazujące 10 najistotniejszych zmiennych dla trzech modeli: *Regresji Logistycznej*, *Lasów Losowych* i *Drzew Wzmacnianych*.

Najważniejszą zmienną dla modelu *Regresji Logistycznej* była Przyczyna\_Upadek (**2.72**). Do zmiennych o bardzo wysokich wartościach należały również:

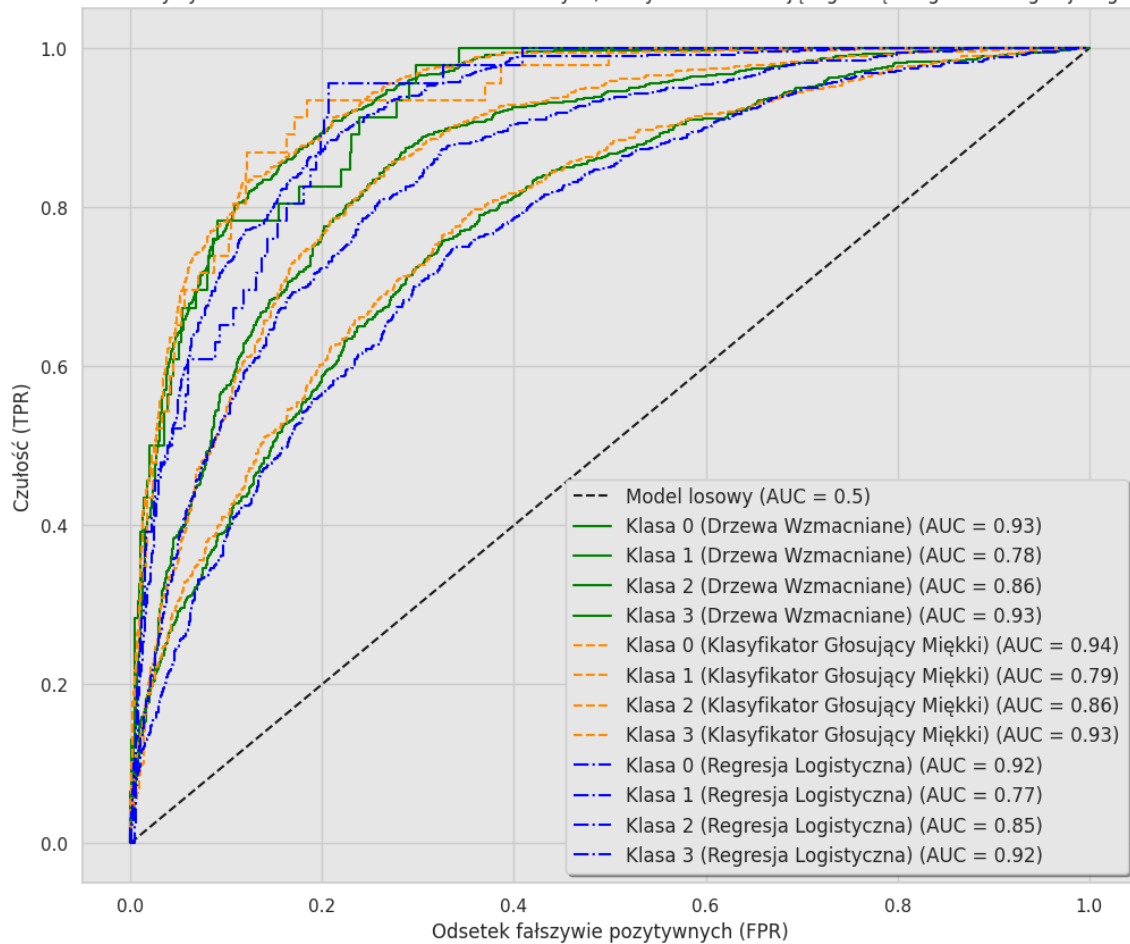
Rodzaj\_wypadku\_Ekstremalne (**2.09**), Rodzaj\_wypadku\_Zbedny\_alarm (**2.03**), Rodzaj\_wypadku\_Prace\_Lesne (**1.96**) oraz Przyczyna\_Brak\_wiedzy/sprzetu (**1.89**). Z 10 najistotniejszych zmiennych, 9 odnosi się do przyczyny oraz rodzaju wypadku, co wskazuje, że są to zmienne ważne i mające duży wpływ na obrażenia w wypadkach górskich.

Najważniejszą zmienną dla modelu *Lasów Losowych* był Czas\_akcji (**0.14**). Zmienne Wiek (**0.10**), Przyczyna\_Upadek (**0.09**), Godzina (**0.08**) oraz Liczba\_rat (**0.07**) również osiągnęły wysokie wartości. Spośród 10 najistotniejszych zmiennych, wiele odnosi się do czasu i charakterystyki wypadku.

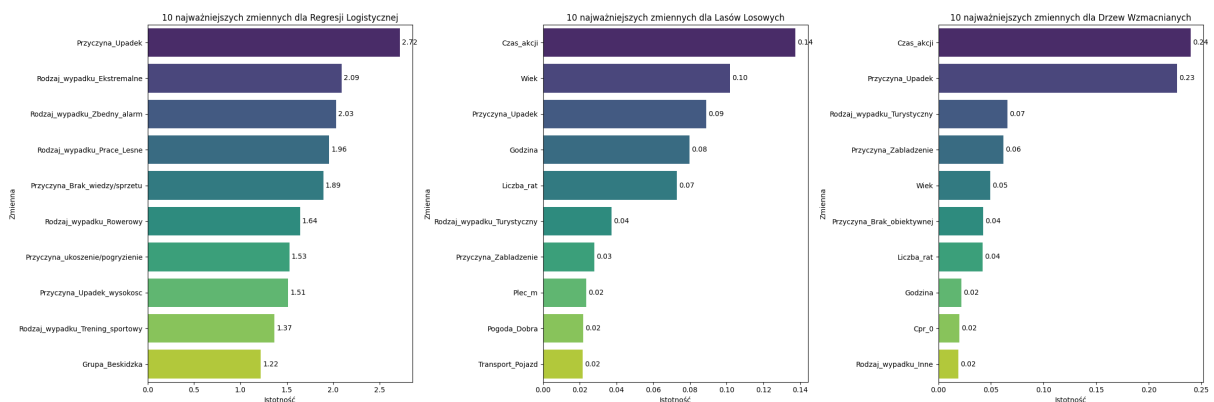
Dla modelu *Drzew Wzmacnianych* najważniejszymi zmiennymi były Czas\_akcji (**0.24**) oraz Przyczyna\_Upadek (**0.23**). Wiele z najistotniejszych zmiennych, odnosi się do przyczyny i rodzaju wypadku.

W celu oceny istotności pierwotnej zmiennej (przed kodowaniem jednoelementowym), zsumowano istotności wszystkich klas powiązanych z daną zmienną. W ten sposób na Rysunku

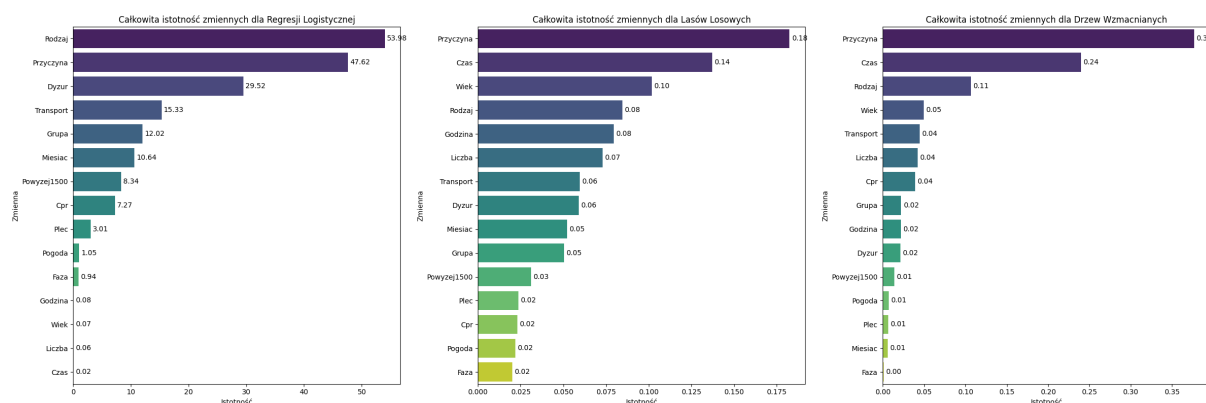
Porównanie krzywych ROC dla modelu Drzew Wzmacnianych, Klasyfikatora Głosującego Miękkiego oraz Regresji Logistycznej



Rysunek 9: Porównanie wykresów krzywych ROC dla modelu Drzew Wzmacnianych, Klasyfikatora Głosującego Miękkiego oraz Regresji Logistycznej



Rysunek 10: 10 najistotniejszych zmiennych dla modelu Regresji Logistycznej, Lasów Losowych oraz Drzew Wzmacnianych



Rysunek 11: Wykresy całkowitej istotności zmiennych dla modelu Regresji Logistycznej, Lasów Losowych oraz Drzew Wzmacnianych

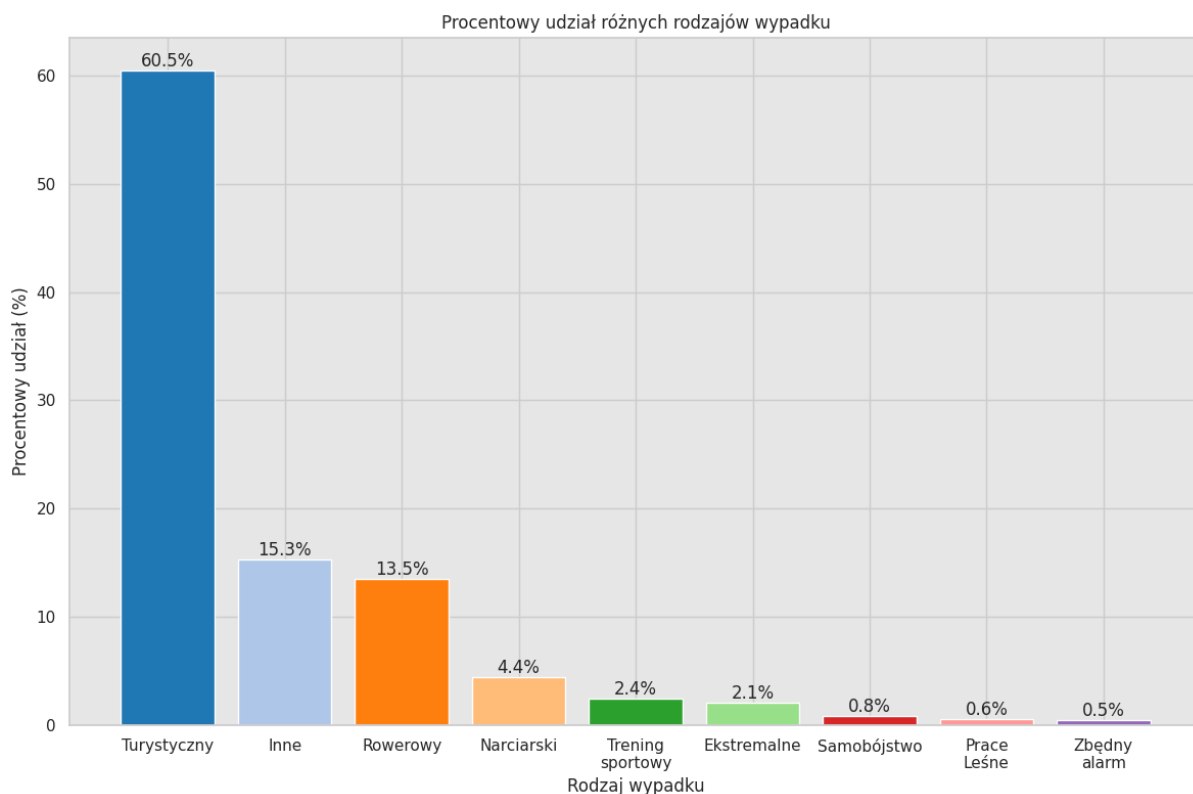
11 przedstawiono wykresy całkowitej istotności zmiennych dla modelu *Regresji Logistycznej*, *Lasów Losowych* oraz *Drzew Wzmacnianych*.

Z rysunku przedstawiającego całkowitą istotność zmiennych dla modelu *Regresji Logistycznej* można zauważyć, że najważniejszymi zmiennymi okazały się Rodzaj\_wypadku (53.98) oraz Przyczyna (47.62), co wskazuje na ich dominujący wpływ na model. Kolejnymi istotnymi zmiennymi są Dyżur (29.52) oraz Transport (15.33). Zmienne Grupa (12.02), Miesiąc (10.64), Powyżej1500 (8.34), Cpr (7.27) także mają znaczący wpływ, podczas gdy pozostałe zmienne, takie jak: Płeć (3.01), Pogoda (1.05), Faza\_tygodnia (0.94), Godzina (0.08), Wiek (0.07), Liczba\_ratowników (0.06) oraz Czas (0.02), mają mniejszy wpływ na model.

Z rysunku przedstawiającego całkowitą istotność zmiennych dla modelu Lasów Losowych wynika, że najważniejszymi zmiennymi są Przyczyna (0.18) oraz Czas (0.14), co wskazuje na ich dominujący wpływ. Kolejne istotne zmienne to Wiek (0.10), Rodzaj\_wypadku (0.08) i Godzina (0.08). Zmienne Liczba\_ratowników (0.07), Transport (0.06), Dyżur (0.06), Miesiąc (0.05) oraz Grupa (0.05) również osiągają wysokie wartości. Pozostałe zmienne, takie jak Powyżej1500 (0.03), Płeć (0.02), Cpr (0.02), Pogoda (0.02) oraz Faza (0.02), osiągają niższe wartości, co sugeruje ich mniejszy wpływ na model.

Z rysunku przedstawiającego całkowitą istotność zmiennych dla modelu Drzew Wzmacnianych wynika, że najistotniejszą zmienną jest Przyczyna (0.38). Kolejne w rankingu istotności są Czas (0.24) oraz Rodzaj\_wypadku (0.11). Powyższe trzy zmienne mają zdecydowanie najwyższy wpływ na model. Istotnymi zmiennymi są również Wiek (0.05), Transport (0.04), Liczba\_ratowników (0.04), Cpr (0.04). Pozostałe zmienne, takie jak Grupa (0.02), Godzina (0.02), Dyżur (0.02), Powyżej1500 (0.01), Pogoda (0.01), Płeć (0.01), Miesiąc (0.01) oraz Faza\_tygodnia (0.00), osiągają niższe wartości, co sugeruje ich ograniczony wpływ na analizę.

Przeprowadzono także analizę najistotniejszych zmiennych biorąc pod uwagę wyniki powyższych trzech modeli.



Rysunek 12: Procentowy udział każdego rodzaju wypadku

Rysunek 12 przedstawia procentowy udział każdego rodzaju wypadku. Zdecydowanie dominującym rodzajem wypadku jest Turystyczny, który stanowi **60.5%** wszystkich wypadków. Kolejne najczęstsze rodzaje to Rowerowy (**13.5%**) oraz Inne (**15.3%**). Wypadki Narciarski stanowią **4.4%**, a wypadki związane z Trening sportowy to **2.4%**. Pozostałe kategorie, takie jak wypadki Ekstremalne, Samobójstwo, Prace Leśne oraz Zbędny alarm, mają znacznie mniejszy udział, nie przekraczający **2.1%**.

Na Wykresie 13 przedstawiono trzy najczęstsze przyczyny każdego rodzaju wypadu.

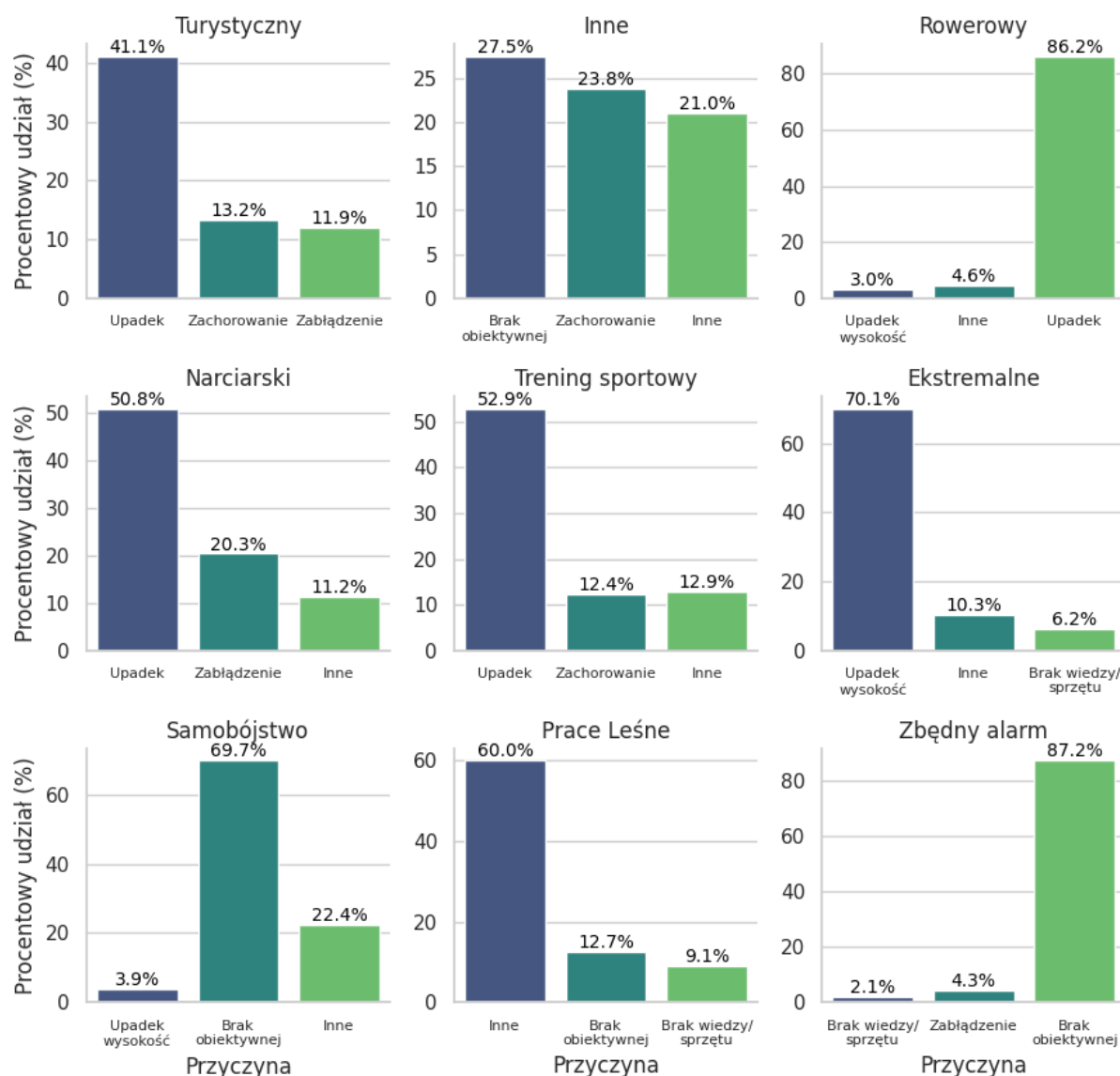
Upadek jest częstą przyczyną w różnych rodzajach wypadku, występując zarówno w wypadkach turystycznych, rowerowych, narciarskich jak i w treningach sportowych. Jest dominującą przyczyną wypadków rowerowych (**86.2%**), narciarskich (**50.8%**) oraz w treningach sportowych (**52.9%**).

Zachorowanie również często pojawia się jako przyczyna wypadków, szczególnie w wypadkach turystycznych (**13.2%**), innych (**23.8%**) oraz w treningach sportowych (**12.4%**).

Istnieją również specyficzne dominujące przyczyny w innych rodzajach wypadku. Na przykład, w przypadku samobójstw, Brak obiektywnej przyczyny stanowi (**69.7%**). W zbędnych alarmach również dominuje Brak obiektywnej przyczyny (**87.2%**). W wypadkach o rodzaju prace leśne, najczęstszą przyczyną jest kategoria Inne (**60.0%**) a głównym powodem wypadków ekstremalnych jest Upadek z wysokości (**70.1%**).

Rysunki te pokazują, że w przypadku większości rodzajów wypadków, dominująca przyczyna stanowi znaczny procent wszystkich przyczyn w tej kategorii.

3 najczęstsze przyczyny każdego rodzaju wypadku



Rysunek 13: 3 najczęstsze przyczyny każdego rodzaju wypadku

Na Rysunku 14 przedstawiono procentowy udział kategorii obrażeń w różnych rodzajach wypadków.

Kategoria 0 (brak obrażeń) ma największy udział w zbędnych alarmach (**97.9%**) oraz znaczący udział w wypadkach turystycznych (**30.6%**) i samobójstwach (**47.4%**).

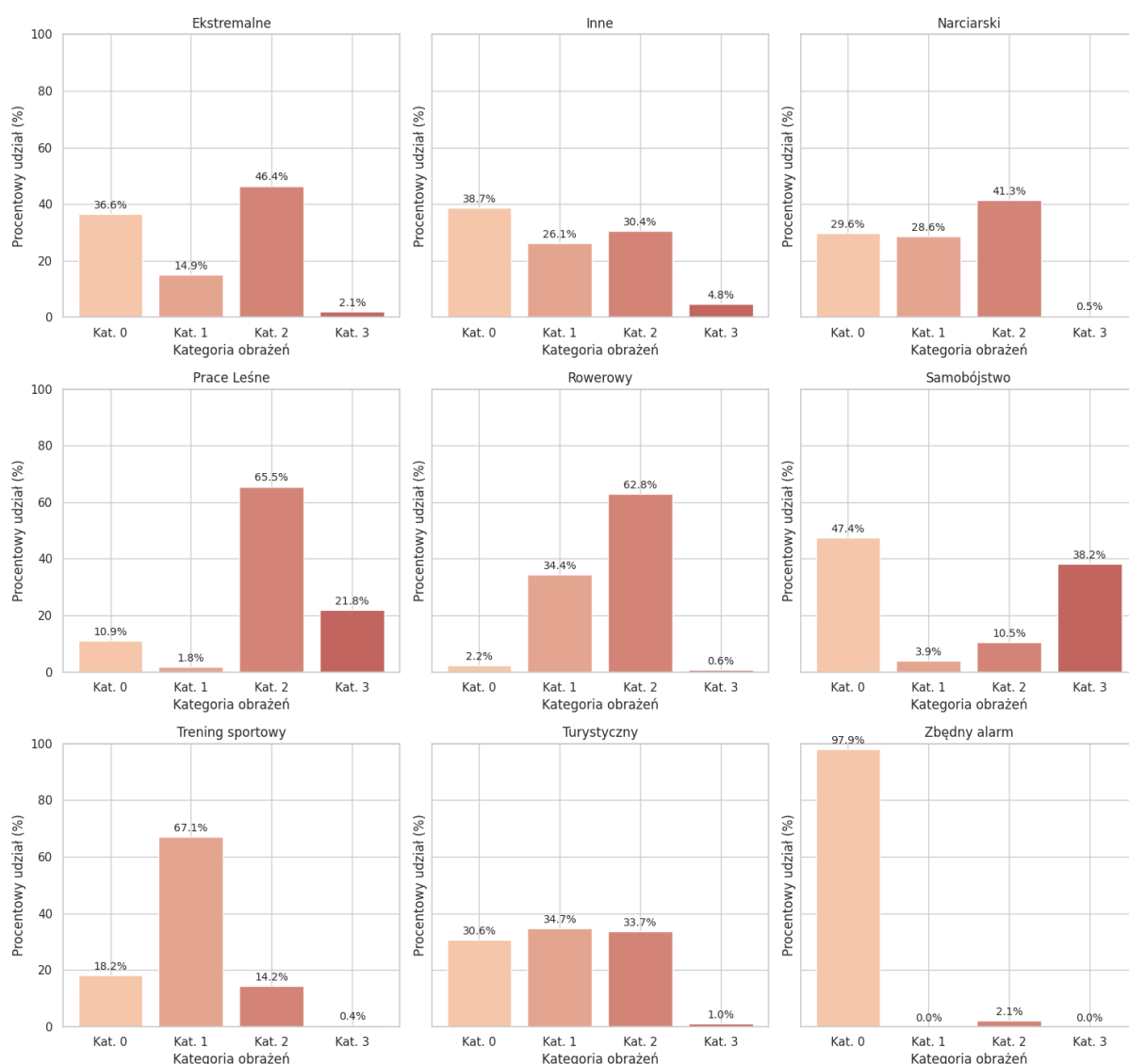
Kategoria 1 (obrażenia lekkie) jest szczególnie dominująca w treningach sportowych (**67.1%**) oraz ma znaczny udział w wypadkach turystycznych (**34.7%**) i innych (**26.1%**).

Kategoria 2 (obrażenia średnie i poważne) dominuje w wypadkach rowerowych (**62.8%**), ekstremalnych (**46.4%**) i w pracach leśnych (**65.5%**).

Kategoria 3 (wypadki śmiertelne) jest dominująca w samobójstwach (**38.2%**), oraz ma znaczny udział w wypadkach związanych z pracami leśnymi (**21.8%**).

Rysunki te pokazują, że w przypadku większości rodzajów wypadków, dominująca kate-

Procentowy udział kategorii obrażeń w każdym rodzaju wypadku



Rysunek 14: Procentowy udział kategorii obrażeń w każdym rodzaju wypadku

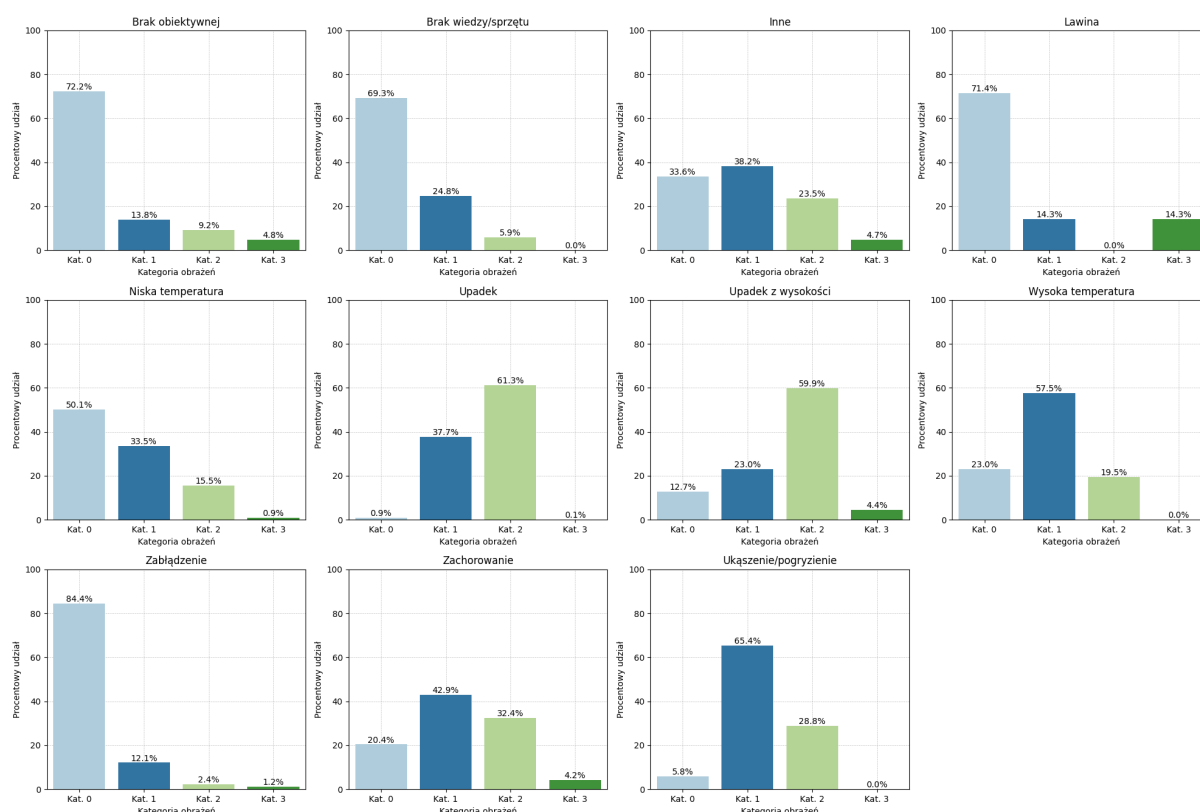
gora obrażeń stanowi znaczny procent wszystkich obrażeń w tej kategorii. Wypadki śmiertelne (kategoria 3) są szczególnie istotne w przypadkach samobójstw (**38.2%**) oraz wypadków w pracach leśnych (**21.8%**), co wskazuje na wysokie ryzyko związane z tymi zdarzeniami.

Na Rysunku 15 przedstawiono procentowy udział kategorii obrażeń w odniesieniu do różnych przyczyn wypadków.

Kategoria 0 (brak obrażeń) ma największy udział w przypadkach zabłądzenia (**84.4%**) oraz znaczący udział w przyczynach takich jak brak obiektywnej przyczyny (**72.2%**) i lawina (**71.4%**).

Kategoria 1 (obrażenia lekkie) jest szczególnie dominująca w przypadkach wysokiej temperatury (**57.5%**) oraz ma znaczny udział w upadkach (**37.7%**) i przyczynach związanych z brakiem wiedzy lub sprzętu (**24.8%**).

Procentowy udział kategorii obrażeń w odniesieniu do przyczyn wypadków



Rysunek 15: Procentowy udział kategorii obrażeń w odniesieniu do różnych przyczyn wypadków

Kategoria 2 (obrażenia średnie i poważne) dominuje w przypadkach upadków (**61.3%**), upadków z wysokości (**59.9%**), zachorowań (**32.4%**) oraz ukąszeń/pogryzień (**28.8%**).

Kategoria 3 (wypadki śmiertelne) ma największy udział w przypadkach związanych z lawinami (**14.3%**). Stosunkowo wysoki udział przypada również w wypadkach, w których nie było obiektywnej przyczyny (**4.8%**) oraz, gdy przyczyną była kategoria inne (**4.7%**), upadek z wysokości (**4.4%**) oraz zachorowanie (**4.2%**).

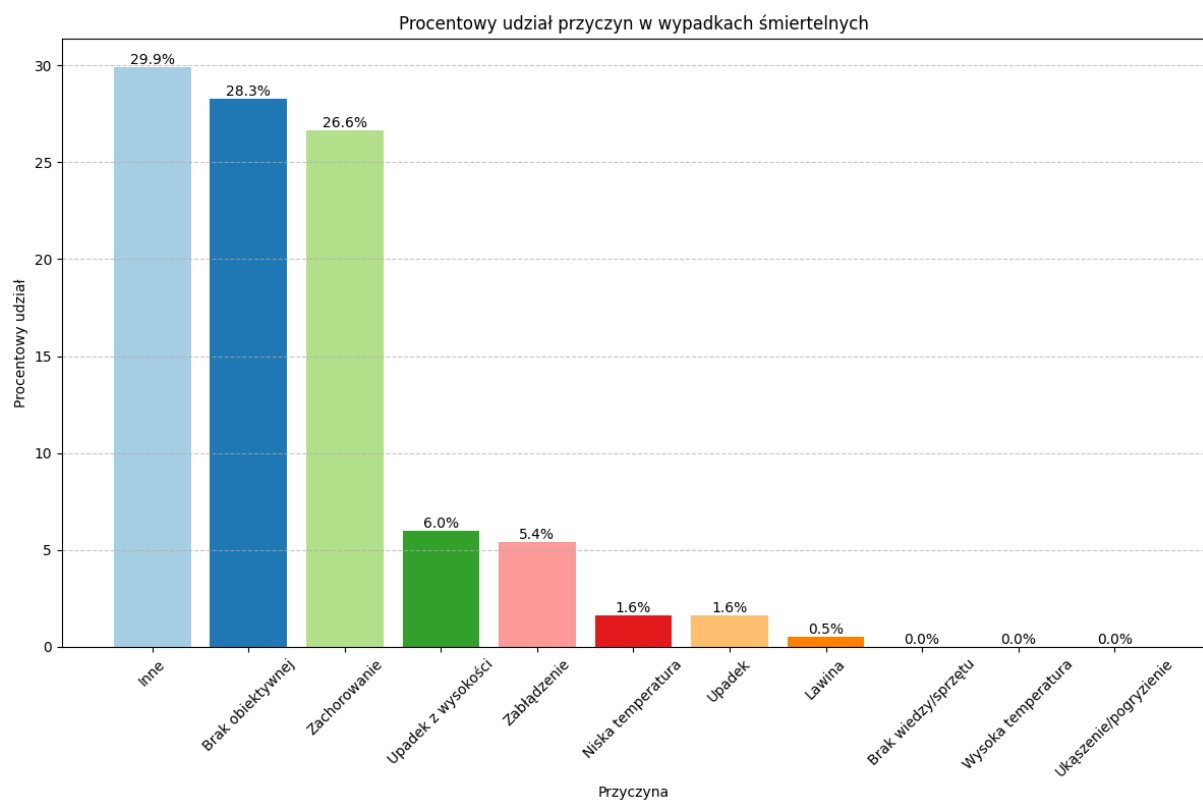
Na Rysunku 16 przedstawiono procentowy udział przyczyn w wypadkach śmiertelnych.

Trzy najczęstsze przyczyny to: Inne (**29.9%**), Brak obiektywnej przyczyny (**28.3%**) oraz Zachorowanie (**26.6%**). Upadek z wysokości (**6.0%**) oraz Zabłądzenie (**5.4%**) również są istotnymi przyczynami wypadków śmiertelnych, natomiast pozostałe przyczyny charakteryzują się małą śmiertelnością, mogąą jednak wynikać z małej liczby wypadków o tej przyczynie.

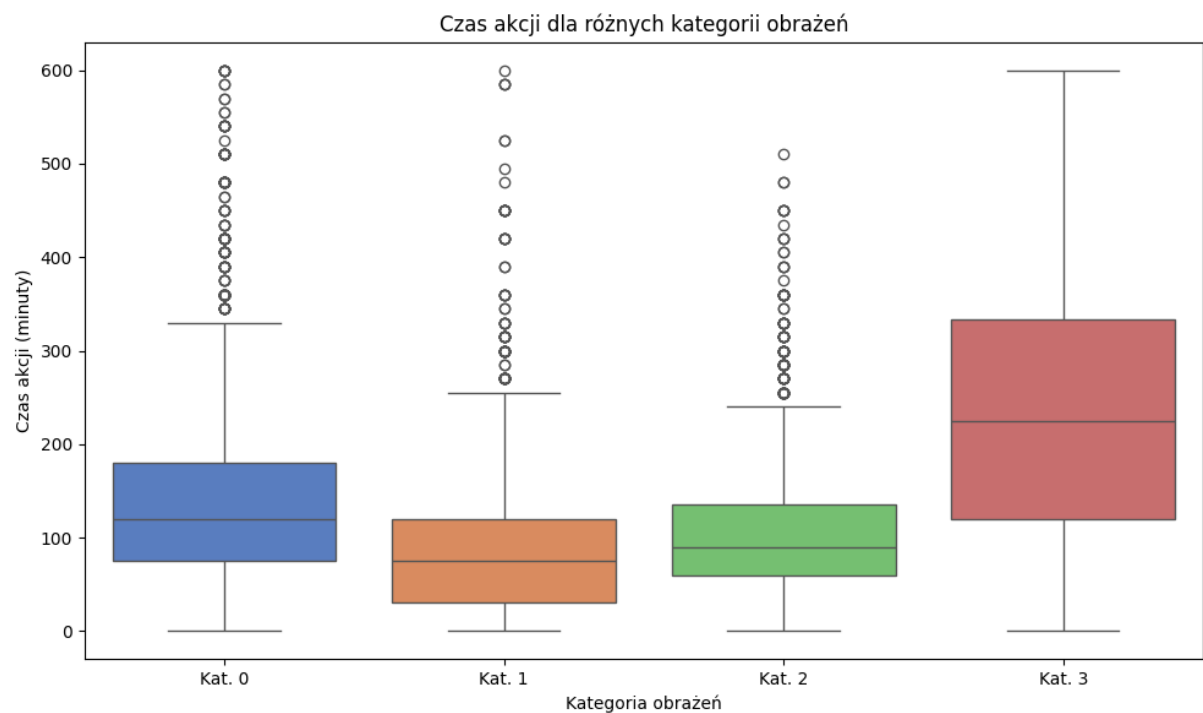
Na Rysunku 17 przedstawiono histogramy czasu akcji w zależności od kategorii obrażeń.

Analiza wykresów wskazuje, że połowa przypadków Kategorii 0 mieści się w przedziale od 80 do 180 minut z medianą na poziomie 120 minut. Zakres czasu działań waha się przeważnie między 0 a 340 minut z wartościami odstającymi dochodzącymi nawet do 600 minut.

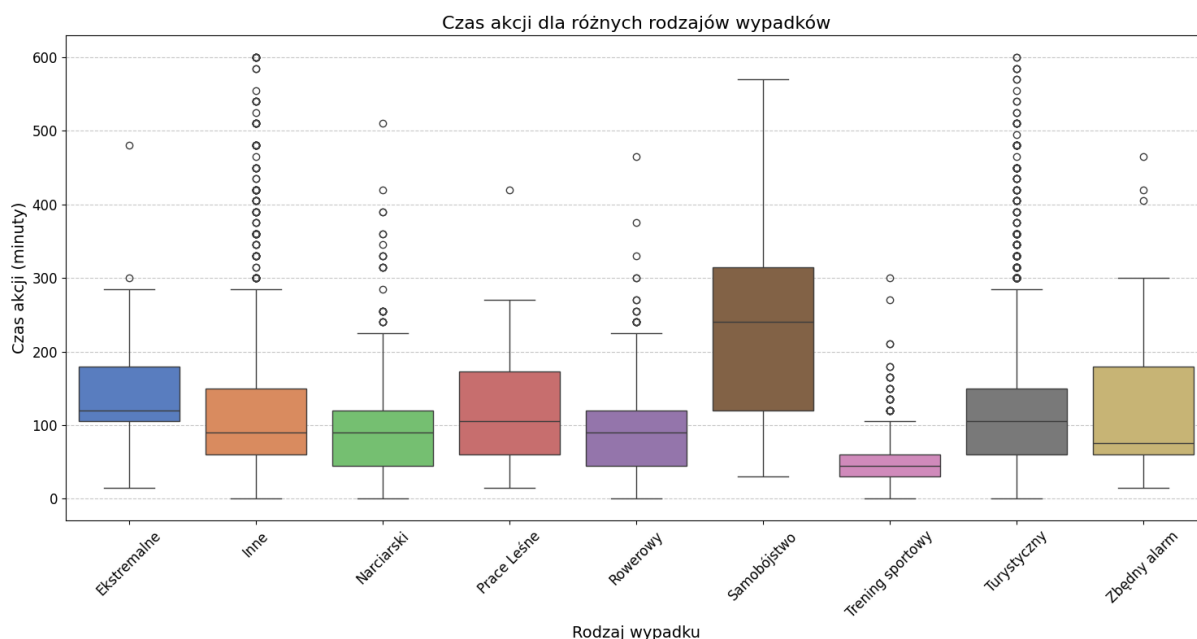




Rysunek 16: Procentowy udział przyczyn w wypadkach śmiertelnych



Rysunek 17: Wykres pudełkowy czasu akcji w zależności od kategorii obrażeń



Rysunek 18: Czas akcji dla różnych rodzajów wypadków

W Kategorii 1 połowa wypadków trwa między 40 a 120 minut. Mediana czasu akcji to około 80 minut. Wypadki z obrażeniami lekkimi charakteryzują się najkrótszym czasem trwania działań wśród wszystkich kategorii. Wartości odstające również dochodzą do 600 minut.

W Kategorii 2 zakres między pierwszym a trzecim kwartylem to 60 - 140 minut, z medianą na poziomie 100 minut. W porównaniu do obrażeń lekkich, obrażenia średnie i ciężkie trwają dłużej, lecz nie jest to duża różnica. Wartości odstające także są wysokie.

W Kategorii 3 dłuższe akcje są bardziej powszechne. Połowa wypadków śmiertelnych charakteryzuje się czasem akcji rozłożonym od 140 do 320 minut, z licznymi przypadkami również w wyższych przedziałach czasowych. Mediana wynosząca 240 minut jest wyraźnie najwyższą spośród wszystkich kategorii. Akcje związane z wypadkami śmiertelnymi są najdłuższe wśród badanych kategorii obrażeń.

Na Rysunku 18 przedstawiono wykres pudełkowy czasu akcji w zależności od rodzaju wypadków.

Wypadki o rodzaju "Samobójstwo", charakteryzują się wyraźnie najszerszym zakresem czasowym działań. Mediana osiągająca wartości około 220 minut, oraz wypadki często trwające między 200 a 300 minut potwierdzają, że ten rodzaj wypadków ma największą tendencję do długotrwałych akcji.

Mediana czasu trwania akcji wynosząca 50 minut, oraz bardzo wąski zakres czasu (25-75 minut) między pierwszym a trzecim kwantylem pozwala określić, że są to wypadki wyraźnie krótsze od pozostałych. Zdarzają się jednak wartości odstające, z maksymalną wartością około 300 minut.

Zbędne alarmy mają nietypowo rozłożony czas akcji. Mediana wynosi około 75 minut,

lecz szeroki zakres czasu trwania akcji pozwala stwierdzić, że zbędne alarmy mają tendencję do długiego czasu trwania.

W przypadku prac leśnych, akcje ratunkowe najczęściej trwają około 100 minut, jednak zakres między pierwszym a trzecim kwantylem jest dość duży sięgający około 3 godzin.

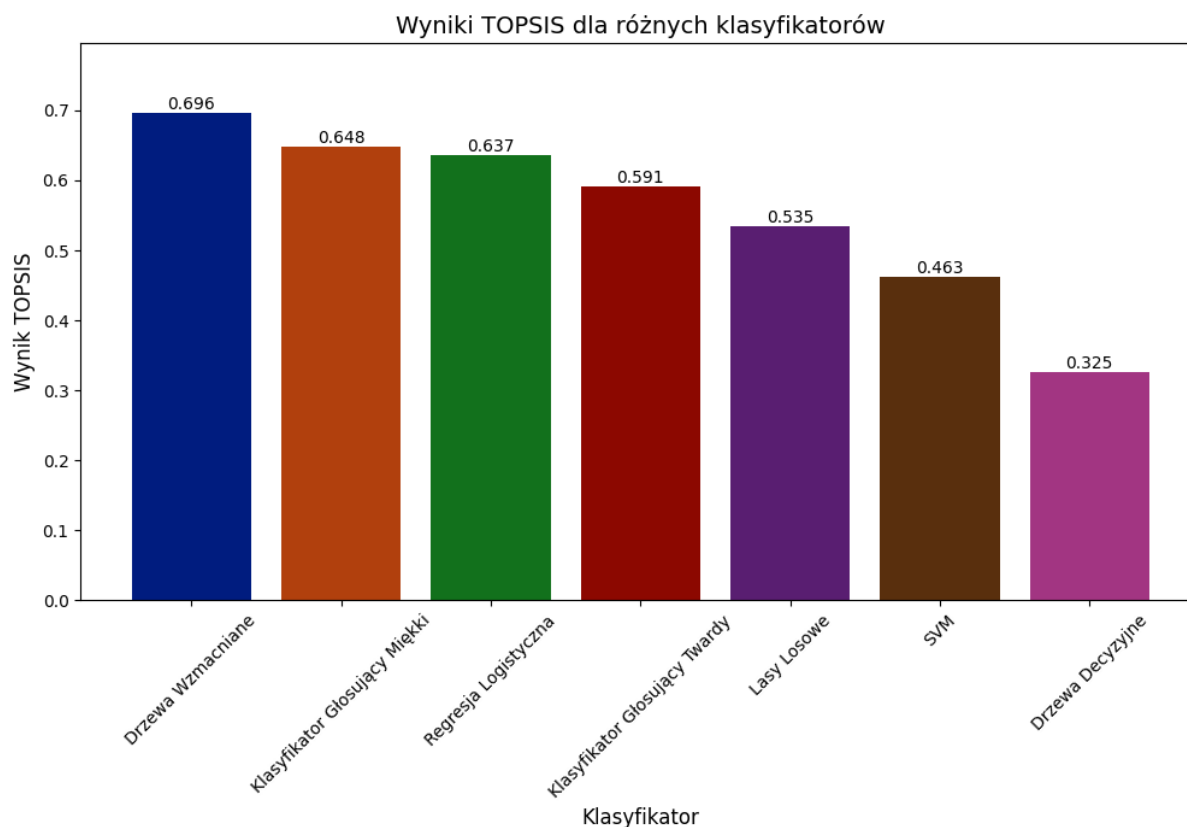
Dla wypadków turystycznych, większość akcji ratunkowych trwa od 75 do 150 minut, z medianą w okolicach 100 minut. Zdarzają się także wartości odstające dochodzące nawet do 600 minut.

Bardzo podobnie wyglądają wypadki o rodzaju "Inne". Posiadają medianę nieco poniżej 100 minut, wartości minimalne, maksymalne oraz kwantyle pierwszy i drugi na tym samym poziomie co wypadki turystyczne, oraz sporo wartości odstających.

Wypadki rowerowe i narciarskie również charakteryzują się dużym podobieństwem w czasie trwania akcji. Mediana wynosi nieco poniżej 100 minut, wartość minimalna i maksymalna są na tym samym poziomie. Zdarzają się także wartości odstające w obu rodzajach wypadków.

Wypadki ekstremalne mają medianę w okolicach 120 minut. Wartości między pierwszym a trzecim kwantylem są większe niż w wypadkach turystycznych, lecz występują znacznie mniej wartości odstających. Można stwierdzić, że trwają one najdłużej po wypadkach z rodzaju "Samobójstwo"

Na Rysunku 19 przedstawiono wyniki zastosowanej metody TOPSIS.



Rysunek 19: Wyniki TOPSIS

Biorąc pod uwagę wszystkie obliczone metryki, ze szczególnym uwzględnieniem poprawnego przewidywania wypadków śmiertelnych, najwyższy wynik TOPSIS uzyskał model *Drzew Wzmacnianych* z wynikiem **0.696**. Pozwala to stwierdzić, że jest to najlepszy klasyfikator dla analizowanych danych. Kolejne w rankingu to *Klasyfikator Głosujący Miękki* (**0.648**) oraz *Regresja Logistyczna* (**0.637**).

Model *Klasyfikator Głosujący Twardy* uzyskał wynik **0.591**, co umieściło go na czwartym miejscu. *Lasy Losowe* osiągnęły wynik **0.535**, a *SVM* uzyskał **0.463**. Najniższy wynik uzyskał model *Drzew Decyzyjnych*, osiągając **0.325**.

## 5. Dyskusja

W procesie przetwarzania wstępnego zastosowano technikę RFECV, a następnie technikę SMOTE. W pracy [53] autorzy szczegółowo analizują wpływ tych technik na skuteczność modeli predykcyjnych w kontekście prognozowania bankructwa. Zastosowanie RFECV pozwala na wyeliminowanie zbędnych zmiennych, co zmniejsza złożoność modelu i poprawia jego interpretowalność, podczas gdy SMOTE pomaga zbalansować nierówne klasy w zbiorze danych, gdyż przypadki bankructwa są zazwyczaj znacznie rzadsze niż przypadki firm funkcjonujących. Badania autorów [53] wykazały, że połączenie tych technik może prowadzić do uzyskania lepszych wyników niż ich stosowanie oddzielnie. Prowadzona przez nich analiza wykazała, że wpływ równoważenia klas jest znaczący w poprawie jakości modeli, natomiast selekcja zmiennych (RFECV) mimo że jest ważna, ma ograniczony wpływ w porównaniu do technik równoważenia klas. Ponadto, w badaniu porównano różne techniki uczenia maszynowego, takie jak Lasy Losowe, AdaBoost, Drzewa Wzmacniane czy Bagging. Stwierdzono, że algorytm AdaBoost wykazuje najslabsze wyniki, natomiast Drzewa Wzmacniane osiągają najwyższą dokładność predykcji.

W przedstawianej pracy rozważano różną kolejność zastosowania techniki RFECV i SMOTE. Najlepszą wydajność modeli uzyskano po zastosowaniu RFECV przed syntetycznym generowaniem próbek. Zastosowana kolejność tych technik znacząco wpływała na wymiarowość danych przekazywanych dalej do modeli. Dzięki selekcji zmiennych, syntetycznie generowane próbki klasy mniejszościowej wykorzystywały tylko najbardziej istotne zmienne, co zapobiegało tworzeniu sztucznej istotności, która mogłaby wpłynąć na RFECV i przyczynić się do nadmiernego dopasowania przy odwrotnej kolejności zastosowanych metod.

Redukcja zmiennych pozwoliła zmniejszyć ilość danych przekazywanych do modelu, co przyspieszyło proces trenowania i dostosowywania parametrów modelu. Kolejnym wnioskiem, potwierdzającym wyniki pracy [53], jest niewielki wpływ dostosowywania parametrów modeli. Różnice w wynikach przeważnie nie dochodziły do jednego punktu procentowego, a w wielu przypadkach zastosowanie modeli z domyślnymi parametrami dawało najlepsze rezultaty.

W prezentowanej pracy najlepsze rezultaty osiągnął model Drzew Wzmacnianych. Podobne wnioski wyciągnięto w pracach [54], [55], [56], w których porównywano różne modele uczenia maszynowego i metod statystycznych, często pokrywających się z modelami zastosowanymi w niniejszej pracy. W tych badaniach również zmagano się z problemem niezbalansowanych klas, predykcją ciężkości obrażeń lub innych medycznych aspektów. Modele Drzew Wzmacnianych i ich modyfikacje uzyskiwały również najlepsze wyniki w pracach związanych z przewidywaniem występowania nowotworów, czy ostrego uszkodzenia nerek [57], [58], [59].

W niektórych badaniach o zbliżonej tematyce najlepszym modelem okazywały się Lasy Losowe [60], mimo że model Drzew Wzmacnianych także był brany pod uwagę. W prezentowanej pracy Lasy Losowe osiągnęły gorsze wyniki, lecz w różnych badaniach metoda ta uzyskuje często bardzo wysokie dokładności.

W pracy [60] wykorzystano różnorodne modele uczenia maszynowego do identyfikacji stopnia ciężkości wypadków na podstawie zmiennych wyodrębnionych z obrazów wypadków. W tej pracy model Lasów Losowych osiągnął najlepszą wydajność, uzyskując wartość AUC na poziomie 0,75. Lasy Losowe uzyskały lepsze wyniki od takich algorytmów jak model Maszyny Wektorów Nośnych, Regresja Logistyczna, Drzewa Decyzyjne, AdaBoost oraz Drzewa Wzmacniane. Wyodrębnione zmienne użyte do trenowania modeli obejmowały kształt (ang. Hu Moments), teksturę (ang. Local Binary Pattern - LBP) oraz gradient intensywności (ang. Histogram of Oriented Gradients - HoG).

Drzewa Wzmacniane mogą skutecznie radzić sobie z danymi o podobnej strukturze, jak te wykorzystane w niniejszej pracy, ponieważ adaptują się do niezbalansowanych danych, nadając większą wagę trudniejszym do klasyfikacji przypadkom. Algorytm iteracyjnie minimalizuje funkcję błędu, co pozwala na stopniowe poprawianie klasyfikacji przypadków należących do mniejszościowej klasy. Dzięki temu model lepiej radzi sobie z niezbalansowanymi klasami i odnosi wysoką skuteczność w przewidywaniu obrażeń, które często mogą być niezbalansowane.

Wyniki przewidywań uzyskane w ramach niniejszej pracy, z najlepszym modelem osiągającym dokładność na poziomie 67%, wskazują na potrzebę dalszej optymalizacji modeli predykcyjnych. Jednym z głównych ograniczeń stworzonych modeli był brak zmiennych medycznych, takich jak wskaźniki fizjologiczne, powiązane choroby, czy nawet lokalizacja urazu, które mogłyby znacząco poprawić ich zdolność predykcyjną. W literaturze wykazano, że te czynniki mają istotny wpływ na dokładność modeli predykcyjnych w ocenie ciężkości obrażeń [61], [62], [63], [64].

W pracy [61], na przykład, autorzy przeprowadzili systematyczny przegląd narzędzi do oceny ciężkości obrażeń, takich jak ISS (ang. Injury Severity Score), NISS (ang. New Injury Severity Score), TRISS (ang. Trauma and Injury Severity Score) i ICISS (ang. International Classification of Diseases-based Injury Severity Score). Celem przeglądu było podsumowanie wyników dotyczących zdolności tych narzędzi do przewidywania śmiertelności pacjentów z urazami oraz zidentyfikowanie źródeł heterogeniczności w badaniach. Wykazano, że skale takie jak NISS i TRISS, które uwzględniają wskaźniki fizjologiczne i lokalizację urazu, osiągają wysoką skuteczność w przewidywaniu śmiertelności i innych wyników klinicznych.

W środowisku naukowym toczy się dyskusja, która ze skal oceny obrażeń daje najlepsze rezultaty w przewidywaniu zgonów oraz innych klas obrażeń, lecz wszystkie brane pod uwagę skale wykorzystują w jakimś stopniu informacje o parametrach życiowych, lokalizacji i rodzaju urazu, a także stanie świadomości poszkodowanego [62], [63], [64].

W pracy [62] przeanalizowano dane ponad dziewięć tysięcy pacjentów pourazowych ocenianych w centrum urazowym poziomu I przez okres 30 miesięcy. Celem było ustalenie, czy PTS (ang. Physiologic Trauma Score), oparty na wyniku SIRS (ang. Systemic Inflammatory Response Syndrome) w połączeniu z innymi wskaźnikami, może dokładnie przewidzieć śmiertelność pacjentów. Wynik SIRS, obliczany na podstawie takich parametrów jak tempera-

tura, tętno, częstość oddechów i liczba neutrofilów, był używany wraz ze wskaźnikami takimi jak skala Glasgow, wiek, płeć oraz rasa w modelach regresji logistycznej do przewidywania ryzyka zgonu. Wyniki pokazały, że model PTS, łączący wynik SIRS ze skalą Glasgow i wiekiem, był równie skuteczny jak bardziej złożone modele, takie jak TRISS, a nawet przewyższał ISS w przewidywaniu śmiertelności. Badanie wykazało, że proste fizjologiczne wskaźniki mogą skutecznie przewidywać wyniki u pacjentów pourazowych, co może ułatwić szybką ocenę stanu pacjenta na miejscu zdarzenia.

W pracy [63] autorzy opracowali i zweryfikowali model prognostyczny oparty na łatwo dostępnych zmiennych klinicznych, takich jak wiek, wynik skali Glasgow, ISS, reakcja źrenic oraz obecność krwiaka w tomografii komputerowej. Wyniki wykazały, że prosty model prognostyczny zawierający te zmienne jest skuteczny w przewidywaniu przeżywalności pacjentów z urazami mózgu, co potwierdzono zarówno na treningowej, jak i na walidacyjnej grupie pacjentów. Model wykazał dobrą zdolność rozróżniania i kalibracji, co wskazuje na jego użyteczność w praktyce klinicznej do szybkiej oceny ryzyka i wspomagania decyzji medycznych.

Dodatkowo, skala NACA, która oryginalnie została zastosowana przez ratowników GOPR do oceny ciężkości obrażeń, niesie pewne kontrowersje. Jak wykazano w badaniach, jej subiektywny charakter może wpływać na dokładność ocen, szczególnie gdy jest stosowana przez mniej doświadczony personel [65], [66].

Celem pracy [65] było zbadanie, czy istnieje związek pomiędzy oceną ciężkości choroby i urazu a doświadczeniem lekarzy medycyny ratunkowej za pomocą skali NACA. Autorzy chcieli sprawdzić, czy doświadczenie medyków wpływa na dokładność oceniania stanu zagrożenia pacjentów. W badaniu [65] wykazano, że medycy z mniejszym doświadczeniem oceniają ciężkość urazów niżej niż ich bardziej doświadczeni koledzy, co może prowadzić do niedoszacowania rzeczywistej ciężkości obrażeń.

Podobnie w badaniu [66] wykazano, że chociaż skala NACA dobrze koreluje z umieralnością i przeniesieniem na oddział intensywnej terapii, jej subiektywny charakter może prowadzić do dużych rozbieżności w ocenie ciężkości urazów. Autorzy sugerują również, że mimo wysokiej trafności prognostycznej skali NACA, warto rozważyć zastosowanie dodatkowych lub alternatywnych skal oceny, aby jeszcze bardziej zwiększyć precyzję diagnozowania i przewidywania wyników klinicznych. Włączenie innych narzędzi, takich jak skala ISS czy RTS (ang. Revised Trauma Score), mogłoby dostarczyć dodatkowych informacji i ułatwić podejmowanie decyzji medycznych. Autorzy podkreślają, że kompleksowe podejście, które łączy różne skale oceny, może poprawić jakość opieki nad pacjentami urazowymi oraz wspierać bardziej precyzyjne i spójne decyzje w krytycznych momentach ratunkowych.

Z ankiety [67] przeprowadzonej w 2022 roku wśród ratowników górskich i medyków wynika, że wielu z nich ma wątpliwości co do oceny poszkodowanych za pomocą skali NACA. Spośród medyków, tylko 31.6% uważa, że ratownicy górscy są w stanie prawidłowo określić obrażenia za pomocą tej skali, a wśród ratowników KKPP (Kurs Kwalifikowanej Pierwszej Pomocy) tylko 11.2% nie ma problemów z jej użyciem. Wyniki te są zgodne z literaturą, która

również poddaje w wątpliwość skuteczność stosowania skali NACA przez osoby bez wykształcenia medycznego [65], [66].

W celu zminimalizowania błędów popełnionych przez ratowników podczas określania skali NACA, w prezentowanej pracy postanowiono zredukować liczbę kategorii obrażeń do 4. Rozważano różny podział kategorii, lecz podział na najprostsze do rozróżnienia kategorie (brak obrażeń - 0, obrażenia lekkie - 1, obrażenia średnie i ciężkie - 2, oraz wypadki śmiertelne - 3) dawał najlepsze rezultaty.

Podsumowując, włączenie zmiennych medycznych, takich jak wskaźniki fizjologiczne i lokalizacja urazu, mogłoby znacząco poprawić zdolność predykcyjną stworzonych modeli. Ponadto, zastosowanie bardziej obiektywnych i kompleksowych skal oceny obrażeń może zwiększyć dokładność ocen i przewidywań.



## 6. Podsumowanie i wnioski

Niniejsze badania mają istotne zastosowanie praktyczne, wspierając zarówno pracę ratowników górskich poprzez poprawę dokładności oceny ciężkości obrażeń, jak i działania profilaktyczne, poprzez identyfikację kluczowych przyczyn wypadków górskich.

Zastosowanie korelacji rang Spearmana oraz metody RFECV, pozwoliło na dobranie odpowiedniego zestawu zmiennych, niosącego najwięcej informacji i jednocześnie ograniczając wymiarowość danych. Z kolei zastosowanie techniki SMOTE pozwoliło na zbalansowanie mniej licznych klas. Wykorzystane metody pozwoliły na stworzenie optymalnego zestawu danych, wykorzystanego później do trenowania i walidacji modeli.

Ze wszystkich stworzonych modeli, najlepszym okazał się model Drzew Wzmacnianych. Podczas oceny modelu szczególną uwagę zwrócono na jakość przewidywania wypadków śmiertelnych. Model Drzew Wzmacnianych osiągnął najlepszy wynik TOPSIS, techniki biorącej pod uwagę metryki takie jak: Dokładność, Czułość, Prezycja, Miara F1 oraz AUC. Dokładność tego modelu osiągnęła 67%. Włączenie do analizy danych medycznych, takich jak wskaźniki fizjologiczne, powiązane choroby, lokalizacja urazu, czy stan świadomości mogłoby znacząco poprawić zdolność predykcyjną modeli. Zastosowanie bardziej obiektywnych skal oceny obrażeń może zwiększyć dokładność ocen i przewidywań.

Najistotniejszymi zmiennymi w stworzonych modelach okazały się: Rodzaj wypadku, Przyczyna i Czas akcji. Najczęściej występującymi rodzajami wypadków były: Turystyczny (60.5%), Inne (15.3%) oraz Rowerowy (13.5%). Główna przyczyna wypadków była zależna od jego rodzaju, lecz upadek stanowił duży procent każdego z nich, a w wielu stanowił główną przyczynę. Dość często dochodziło do zdarzenia bez obiektywnej przyczyny, oraz na skutek zachorowania. Do najpoważniejszych obrażeń dochodziło podczas prac leśnych oraz wypadków rowerowych. Rodzaj "Samobójstwo" oraz "Prace Leśne" charakteryzowały się najwyższym współczynnikiem zgonów (odpowiednio 38.2% oraz 21.8%). Zabłędzenie, lawina, brak wiedzy/sprzętu, brak obiektywnej przyczyny oraz niska i wysoka temperatura najczęściej prowadziły do braku obrażeń lub obrażeń lekkich. Upadek oraz upadek z wysokości najczęściej skutkował obrażeniami średnimi i ciężkimi. Do tego rodzaju obrażeń dochodziło również często w wypadkach na skutek zachorowania, ukąszenia/pogryzienia, wysokiej temperatury oraz innych. Najwyższą śmiertelnością charakteryzowały się przyczyny: lawina, brak obiektywnej przyczyny, inne, upadek z wysokości oraz zachorowanie. Wypadki śmiertelne charakteryzują się także długim czasem trwania akcji. Prace leśne oraz samobójstwa zajmowały ratownikom najwięcej czasu, co może być wynikiem śmiertelności i powagi obrażeń powiązanych z tymi zdarzeniami. Wypadki turystyczne i ekstremalne najczęściej miały podobny czas trwania, lecz wypadki ekstremalne miały większą tendencję do dłuższych działań. Może to wynikać z sytuacji, w której poszkodowany był w trudno dostępnym terenie lub wymagał użycia technik linowych do ewakuacji, które zajmują więcej czasu. Mimo, że zbędne alarmy najczęściej nie wymagały długiego czasu akcji, to miały tendencję to przedłużania się. Może to być spowo-

dowane częstym prowadzeniem poszukiwań podczas tego rodzaju wypadków, które kończą się odwołaniem akcji na skutek np. zgłoszenia się telefonicznie poszkodowanego lub rodziny. Ewidentnie najkrótszym czasem trwania akcji charakteryzują się treningi sportowe. Być może częste występowanie braku obrażeń oraz obrażeń lekkich, oraz niedaleka odległość od dostępnych dla pojazdów dróg odgrywa tutaj znaczącą rolę.

W celu poprawy jakości udzielanej pomocy, ratownicy podczas treningu powinni skupiać się na najczęstszych przyczynach każdego rodzaju wypadków. Symulacje bardzo poważnych obrażeń powinny trwać dłużej niż przeciętny czas akcji. Ratownicy powinni skupiać się także na edukacji w zakresie rozpoznawania i radzenia sobie w sytuacji zachorowania poszkodowanego. Stworzenie narzędzia umożliwiającego stały i bezpośredni kontakt ze specjalistą w tej dziedzinie mogłoby znacząco poprawić jakość udzielanej pomocy. W celu zmniejszenia liczby wypadków związanych z brakiem sprzętu lub wiedzy, ratownicy powinni przedsięwziąć więcej środków profilaktycznych w postaci edukacji osób zewnętrznych.

Ujednolicony proces zbierania wywiadu medycznego oraz parametrów życiowych i wskaźników fizjologicznych poszkodowanego, mogłoby przyczynić się do dokładniejszych analiz przyczyn i skutków wypadków, a także pomóc ratownikom w ocenie ciężkości obrażeń osób ulegających wypadkom.

## **Literatura**

- [1] Yan M, Shen Y. Traffic accident severity prediction based on random forest. Sustainability. 2022;14(3):1729. Publisher: MDPI.
- [2] Yassin SS, Pooja. Road accident prediction and model interpretation using a hybrid K-means and random forest algorithm approach. SN Applied Sciences. 2020;2.
- [3] Li L, Prato CG, Wang Y. Ranking contributors to traffic crashes on mountainous freeways from an incomplete dataset: A sequential approach of multivariate imputation by chained equations and random forest classifier. Accident Analysis & Prevention. 2020;146:105744.
- [4] Khanum H, Garg A, Faheem MI. Accident severity prediction modeling for road safety using random forest algorithm: an analysis of Indian highways. F1000Research. 2023.
- [5] Delen D, Sharda R, Bessonov M. Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. Accident Analysis & Prevention. 2006;38(3):434-44. Publisher: Elsevier.
- [6] Ustawa z dnia 15 kwietnia 2011 r. o działalności leczniczej; 2011. Published: Dz.U. 2011 nr 112 poz. 654.
- [7] Ratunkowe GOP. O GOPR; 2016. Publication Title: Strona główna. Available from: <https://www.gopr.pl/gopr/o-gopr>.

- [8] Kaganek K, Wróbel M. Analiza wypadkowości turystycznej w Beskidzie Żywieckim w latach 2006–2010. *Kultura Bezpieczeństwa Nauka–Praktyka–Refleksje*. 2016;23:111-30. Publisher: Wyższa Szkoła Bezpieczeństwa Publicznego i Indywidualnego “Apeiron” w Krakowie.
- [9] Bera M, Zaremba A. Baza noclegowa i towarzysząca–wybrane elementy infrastruktury polskich parków narodowych w latach 2005–2015. *Przegląd Turystyczny*.
- [10] Nithya B, Ilango V. Predictive analytics in health care using machine learning tools and techniques. In: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE; 2017. p. 492-9.
- [11] McCarthy RV, McCarthy MM, Ceccucci W. Predictive models using regression. *Applying Predictive Analytics: Finding Value in Data*. 2022:87-121. Publisher: Springer.
- [12] Lee EK. Large-scale optimization-based classification models in medicine and biology. *Annals of biomedical engineering*. 2007;35:1095-109. Publisher: Springer.
- [13] Cuenca LG, Puertas E, Aliane N, Andres JF. Traffic accidents classification and injury severity prediction. In: 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICITE). IEEE; 2018. p. 52-7.
- [14] Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupuleti S, et al. COVID-19 patient health prediction using boosted random forest algorithm. *Frontiers in public health*. 2020;8:357. Publisher: Frontiers Media SA.
- [15] Ding J, Dai Q, Fan W, Lu M, Zhang Y, Han S, et al. Impacts of meteorology and precursor emission change on O<sub>3</sub> variation in Tianjin, China from 2015 to 2021. *Journal of Environmental Sciences*. 2023;126:506-16. Publisher: Elsevier.
- [16] Schonlau M, Zou RY. The random forest algorithm for statistical learning. *The Stata Journal*. 2020;20(1):3-29. Publisher: SAGE Publications Sage CA: Los Angeles, CA.
- [17] Zhao H, Li X, Cheng H, Zhang J, Wang Q, Zhu H. Deep learning-based prediction of traffic accidents risk for Internet of vehicles. *China Communications*. 2022;19(2):214-24. Publisher: IEEE.
- [18] Koma H, Harada T, Yoshizawa A, Iwasaki H. Detecting cognitive distraction using random forest by considering eye movement type. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*. 2017;11(1):16-28. Publisher: IGI Global.
- [19] Zhang Y, Li H, Ren G. Estimating heterogeneous treatment effects in road safety analysis using generalized random forests. *Accident Analysis & Prevention*. 2022;165:106507. Publisher: Elsevier.

- [20] Lee TC, Shah NU, Haack A, Baxter SL. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review. In: Informatics. vol. 7. MDPI; 2020. p. 25. Issue: 3.
- [21] Weinger MB, Slagle J, Jain S, Ordonez N. Retrospective data collection and analytical techniques for patient safety studies. *Journal of biomedical informatics*. 2003;36(1-2):106-19. Publisher: Elsevier.
- [22] Mieronkoski R, Syrjälä E, Jiang M, Rahmani A, Pahikkala T, Liljeberg P, et al. Developing a pain intensity prediction model using facial expression: A feasibility study with electromyography. *PloS one*. 2020;15(7):e0235545. Publisher: Public Library of Science San Francisco, CA USA.
- [23] Yannis G, Dragomanovits A, Laiou A, Richter T, Ruhl S, Torre FL, et al. Use of accident prediction models in road safety management—an international inquiry. *Transportation research procedia*. 2016;14:4257-66. Publisher: Elsevier.
- [24] Adefabi A, Olisah S, Obunadike C, Oyetubo O, Taiwo E, Tella E. Predicting Accident Severity: An Analysis Of Factors Affecting Accident Severity Using Random Forest Model. *ArXiv*. 2023;abs/2310.05840.
- [25] Ratunkowe GOP. Tekst jednolity Statutu Gopr obowiązujący od 21.10.2022, obejmujący zmiany NZD z 04.06.2022; 2022. Available from: <https://www.gopr.pl/assets/goprMain/media/files/66cddcc3-992f-473a-adf8-02e54d046127/gopr-tekst-jednolity-statutu-gopr-obowia-zuja-cy-od-20221021-obejmuja-cy-zmiany-nzd-z-20220604.pdf>.
- [26] Weiss M, Bernoulli L, Zollinger A. The NACA scale. Construct and predictive validity of the NACA scale for prehospital severity rating in trauma patients. *Der Anaesthesist*. 2001;50(3):150-4.
- [27] Fisher RA. Statistical methods for research workers. In: Breakthroughs in statistics: Methodology and distribution. Springer; 1970. p. 66-70.
- [28] Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*. 2002;16:321-57.
- [29] Lemaitre G, et al. imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning; 2017. Publication Title: *Journal of Machine Learning Research*. Available from: <https://jmlr.org/papers/volume18/16-365/16-365.pdf>.
- [30] Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. *Machine learning*. 2002;46:389-422. Publisher: Springer.

- [31] Berrar D, al e. Cross-validation; 2019.
- [32] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*. 1995;14:1137-45. Publisher: Montreal, Canada.
- [33] Pedregosa F, al e. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-30.
- [34] McCullagh P. Generalized linear models. Routledge; 2019.
- [35] Jr DWH, Lemeshow S, Sturdivant RX. Applied logistic regression. John Wiley & Sons; 2013.
- [36] Kleinbaum DG, al e. Logistic regression. Springer; 2002.
- [37] Menard S. Applied logistic regression analysis. 106. Sage; 2002.
- [38] Agresti A. Categorical data analysis. vol. 792. John Wiley & Sons; 2012.
- [39] Breiman L, Friedman J, Olshen R, Stone C. Cart. Classification and regression trees. 1984. Publisher: Wadsworth and Brooks/Cole Monterey, CA, USA.
- [40] Raileanu LE, Stoffel K. Theoretical comparison between the gini index and information gain criteria. *Annals of Mathematics and Artificial Intelligence*. 2004;41:77-93. Publisher: Springer.
- [41] Breiman L. Random forests. *Machine learning*. 2001;45:5-32. Publisher: Springer.
- [42] Ho TK. Random decision forests. In: *Proceedings of 3rd international conference on document analysis and recognition*. vol. 1. IEEE; 1995. p. 278-82.
- [43] Genuer R. Variance reduction in purely random forests. *Journal of Nonparametric Statistics*. 2012;24:543-62.
- [44] Wager S. Asymptotic Theory for Random Forests. *arXiv: Statistics Theory*. 2014.
- [45] Friedman JH. Greedy function approximation: a gradient boosting machine. *Annals of statistics*. 2001:1189-232. Publisher: JSTOR.
- [46] Bishop CM, Nasrabadi NM. Pattern recognition and machine learning. vol. 4. Springer; 2006.
- [47] Géron A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. "O'Reilly Media, Inc."; 2022.
- [48] Vapnik VN, al e. Statistical learning theory. Wiley New York; 1998. Publication Title: Statistical Learning Journal.

- [49] Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995;20:273-97. Publisher: Springer.
- [50] Hastie T, et al. *The elements of statistical learning: data mining, inference, and prediction*. vol. 2. Springer; 2009.
- [51] Madanchian M. A comprehensive guide to the TOPSIS method for multi-criteria decision making. *Sustainable Social Development*. 2023;1(1):2220.
- [52] Roberts RA, Mullis CT. *Digital signal processing*. Addison-Wesley Longman Publishing Co., Inc.; 1987.
- [53] Shahana T, Lavanya V, Bhat AR. Ensemble classifiers for bankruptcy prediction using SMOTE and RFECV. *International Journal of Enterprise Network Management*. 2024;15(1):109-32. Publisher: Inderscience Publishers (IEL).
- [54] Kashifi MT, Ahmad I. Efficient Histogram-Based Gradient Boosting Approach for Accident Severity Prediction With Multisource Data. *Transportation Research Record*. 2022;2676:236-58.
- [55] Tang J, Liang J, Han C, Li Z, Huang H. Crash injury severity analysis using a two-layer Stacking framework. *Accident; Analysis and Prevention*. 2019;122:226-38.
- [56] Liu L, Zhang X, Liu Y, Zhu W, Zhao B. An Ensemble of Multiple Boosting Methods Based on Classifier-Specific Soft Voting for Intelligent Vehicle Crash Injury Severity Prediction. *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*. 2020:17-24.
- [57] Faisal MI, Bashir S, Khan ZS, Khan FH. An Evaluation of Machine Learning Classifiers and Ensembles for Early Stage Prediction of Lung Cancer. *2018 3rd International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST)*. 2018:1-4.
- [58] Lee HC, Yoon H, Nam K, Cho Y, Kim TK, Kim W, et al. Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery. *Journal of Clinical Medicine*. 2018;7.
- [59] Lee HC, Yoon S, Yang SM, Kim W, Ryu H, Jung C, et al. Prediction of Acute Kidney Injury after Liver Transplantation: Machine Learning Approaches vs. Logistic Regression Model. *Journal of Clinical Medicine*. 2018;7.
- [60] Princess BP, Joyce P, Silas S, Rajsingh EB. Machine Learning Approach for Identification of Accident Severity from Accident Images Using Hybrid Features. In: *2020 International Conference for Emerging Technology (INCET)*; 2020. p. 1-4.

- [61] Tohira H, Jacobs I, Mountain D, Gibson N, Yeo A. Systematic review of predictive performance of injury severity scoring tools. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*. 2012;20:63.
- [62] Kuhls D, Malone D, McCarter R, Napolitano L. Predictors of mortality in adult trauma patients: the physiologic trauma score is equivalent to the Trauma and Injury Severity Score. *Journal of the American College of Surgeons*. 2002;194(6):695-704.
- [63] Signorini D, Andrews P, Jones PA, Wardlaw J, Miller J. Predicting survival using simple clinical variables: a case study in traumatic brain injury. *Journal of Neurology, Neurosurgery & Psychiatry*. 1999;66:20-5.
- [64] Schluter PJ. The trauma and injury severity score (TRISS) revised. *Injury*. 2011;42(1):90-6.
- [65] Knapp J, Bernhard M, Hainer C, Sikinger M, Brenner T, Schlechtriemen T, et al. Is there an association between the rating of illness and injury severity and the experience of emergency medical physicians? *Der Anaesthesist*. 2008;57(11):1069-74.
- [66] Weiss M, Bernoulli L, Zollinger A. The NACA scale. Construct and predictive validity of the NACA scale for prehospital severity rating in trauma patients. *Der Anaesthesist*. 2001;50(3):150-4.
- [67] Beżyk B. Prawidłowe wykorzystanie skali NACA przez ratowników górskich; 2022. Available from: <https://t.ly/j7IG>.