

2. MODELE KLASYFIKACYJNE I GRUPOWANIE POJĘCIOWE JAKO NARZĘDZIA WSPOMAGAJĄCE BADANIE WYPADKOWOŚCI W POLSKIM BUDOWNICTWIE

Wojciech Drozd¹

2.1. WPROWADZENIE

Rozproszony charakter działalności budowlanej i często pojedyncze wypadki na odległych od siebie terenach budów nie przyciągają uwagi mediów i tym samym nie są powszechnie znane. Nie znaczy to, że problem zapewnienia bezpiecznej pracy na budowach został pomyślnie rozwiązany. Wprost przeciwnie – budownictwo zajmuje wysokie miejsce w statystykach wypadkowości w Polsce, jest działem gospodarki wysoce wypadkogennym. Duża liczba wypadków w budownictwie i ich śmiertelnych ofiar sprawia, że problem bezpieczeństwa pracy na terenie budowy ma szczególnie duże znaczenie i wymaga nieustannej troski (Hoła, 2016; Drozd, 2016). Jest bardzo aktualny zwłaszcza w sytuacji, gdy zamawiający usługi budowlane stawiają coraz wyższe wymagania dotyczące czasu i kosztu wykonywanych robót przy jednoczesnym zapewnieniu odpowiednio wysokich standardów jakościowych (Hoła i in., 2015). Dla przedsiębiorstw budowlanych oznacza to pośpiech, cięcie kosztów, co niestety często odbywa się kosztem bezpieczeństwa pracy (Drozd, 2017).

Przedstawione w niniejszym rozdziale badania na temat bezpieczeństwa realizacji robót budowlanych przeprowadzono na podstawie danych zawartych w rejestrze prowadzonym przez Okręgowy Inspektorat Pracy w Krakowie. Koncen-

¹Politechnika Krakowska, Wydział Inżynierii Lądowej, e-mail: wdroid@ztob.pk.edu.pl

trują się one na przedstawieniu zależności przyczynowo – skutkowych zagrożeń w budownictwie. Badaniami objęto odnotowane wypadki przy pracy, na budowach Małopolski, w okresie 2014–2016. Jako narzędzia wspomagające analizy zastosowano klasyfikatory statystyczne, w postaci drzew decyzyjnych oraz grupowanie pojęciowe, w postaci analizy skupień. Takie podejście metodologiczne stanowi propozycję powiększenia zbioru wiedzy w zakresie rozwoju naukowych metod oceny zagrożeń na budowach oraz w zakresie możliwości wykorzystania ich w praktyce. Zaproponowane metody badawcze i przedmiot badań kontynuują dotychczasowe osiągnięcia w zakresie podjętego tematu (Choudhry i in., 2008, Shapira i in., 2009, Törner i in., 2009) i jednocześnie zmniejszają lukę wśród tego typu analiz w naszym kraju.

2.2. WYBÓR ZMIENNYCH DO ANALIZ

Przedmiotem analiz w niniejszym rozdziale są dane oportunistyczne (zbierane długotrwale, bez stałych zasad, dostosowane do chwilowych okoliczności, o znaczącej liczebności) o wypadkach budowlanych, przy realizacji robót. Są to dane rzeczywiste, rejestrowane w urzędowych raportach, regularnie uzupełniane, stanowiące źródło informacji do zaawansowanych prac analitycznych.

W Polsce zdarzenia wypadkowe analizowane są przede wszystkim przez urzędy statystyczne oraz inspektoraty pracy, które je ewidencjonują. W niniejszym rozdziale podjęto próbę identyfikacji i analizy zagrożeń na terenie budów w oparciu o karty wypadków (65 szt.), sporządzanych przez inspektorów Okręgowego Inspektoratu Pracy w Krakowie.

2.2.1. STRUKTURA I PORZĄDKOWANIE DANYCH

Struktura gromadzonych zasobów informacyjnych kształtuje się według obowiązujących w inspektoratach pracy szablonów. Sposób uporządkowania danych dla potrzeb niniejszych badań przedstawiono w tabeli 1.

W analizach raportów z wypadków budowlanych uwzględniono także kody warunków do zaistnienia wypadków (tabela 2) jako wielokrotne odpowiedzi dla zmiennych z raportów o wypadkach budowlanych (tabela 1).

Pozyskane dane z raportów zapisano w arkuszu: „Wypadki” (tabela 3a do 3c – ograniczono się do pokazania 2 z 65 obserwacji). Liczby porządkowe w tabeli są tożsame z kolejnymi raportami z wypadków budowlanych, natomiast kolumny przedstawiają wartości zmiennych niezależnych z tabeli 1 oraz wystąpienie zakodowanych warunków do zaistnienia wypadków budowlanych z tabeli 2 (wartość 1 oznacza wystąpienie danego warunku).

Porządkowanie danych związane było z „obróbką” danych, mającą na celu skorygowanie struktury zasobów danych, usunięcie nieprawidłowości, rozwiązanie ewentualnych niejasności i wątpliwości oraz odpowiednie ich przekształcenie

Tabela 1.

Dane o wypadkach budowlanych (zmiennie niezależne). Źródło: opracowanie własne.

Cmc	– czynnik materialny związany z czynnością wykonywaną przez poszkodowanego w chwili wypadku
Cmo	– czynnik materialny związany z odchyleniem od prawidłowości
Cmu	– czynnik materialny będący źródłem urazu
Cw	– czynność wykonywana przez poszkodowanego w chwili wypadku
Ldn	– liczba dni niezdolności do pracy (dni)
Mw	– miejsce powstania wypadku (BUDOWA, OBIEKT REMONTOWANY)
Pp	– proces pracy (RZ – ziemne, TYNK – tynkowanie, MUR – murowe, PRZEB – przebudowa, MON – montaż, DESK – deskowanie, DEM – demontaż, DACH – dachy, INST – instalacyjne, TRANSP – transport, BET – betonowe)
Stp	– staż pracy poszkodowanego (lata)
Sw	– skutek wypadku (lekki, ciężki, śmiertelny)
Sz	– status zatrudnienia (UPO – umowa na czas określony, UPN – umowa na czas nieokreślony, UZ – umowa zlecenie, UD – umowa o dzieło, DG – działalność gospodarcza)
Tw	– typ wypadku (pojedynczy, zbiorowy)
Wdo	– wydarzenie będące odchyleniem od stanu normalnego
Wi	– wiek poszkodowanego (lata)
Wyu	– wydarzenie powodujące uraz
Zwd	– zawód wykonywany (CIEŚLA-STOLARZ, DEKARZ, ELEKTRYK, FLIZIARZ-KAMIENIARZ, GEOLOG, KIEROWCA WÓZKA WIDŁOWEGO, MALARZ BUDOWLANY, MONTER INSTALACJI GAZOWEJ, MONTER INSTALACJI WOD.-KAN., MURARZ-TYNKARZ, OPERATOR KOPARKI, OPERATOR ŻURAWIA WIEŻOWEGO, PRACOWNIK OGÓLNOBUDOWLANY, SPAWACZ, ZBROJARZ-BETONIARZ)

Tabela 2.

Zakodowane wartości zmiennych opisujących warunki do zaistnienia wypadków budowlanych (odpowiedzi dla zmiennych z tabeli 1). Źródło: opracowanie własne.

BK	– brak wymaganych kwalifikacji, (bez badań lekarskich, instruktażu)
BŚI	– brak środków ochrony indywidualnej (niewłaściwe posługiwanie się środkami ochrony indywidualnej)
BŚZ	– brak środków ochrony zbiorowej
DEM	– demontaż elementów obiektu
DP	– drabina przenośna
EB	– element budynku
EIg	– element instalacji gazowej
EN	– elektronarzędzie,
KN	– kontakt z narzędziem
ŁO	– łopata
MB	– materiały budowlane
NCM	– niewłaściwy stan czynnika materialnego

Tabela 2 [cd.]

NPD	– nieodpowiednie przejścia i dojścia
OD	– odłamki
OE	– przecięcie i skaleczenie
OMS	– przeciążenie układu mięśniowo – szkieletowego
PNN	– praca narzędziem niezmechanizowanym
PNO	– wykonywanie prac z polecenia, nie wchodzących w zakres obowiązków
PNZ	– praca narzędziem zmechanizowanym
PODN	– transport ręczny (podnoszenie, opuszczanie)
POR	– porażenie prądem
POŚ	– potknięcie i poślizg
PP	– powierzchnia pozioma (grunt)
PRZ	– przygniecenie
PUZ	– wykonywanie prac bez usunięcia zagrożeń
RM	– ruchome maszyny i urządzenia
RU	– rusztowanie
SCH	– przemieszczanie się (schodzenie, wchodzenie, przechodzenie)
SM	– sposób magazynowania
SP	– spadający przedmiot
SPF	– niewłaściwy stan psycho-fizyczny (spożycie alkoholu, zmęczenie)
SZ	– nieprawidłowe, samowolne zachowanie się pracownika, (brak nadzoru, brak koncentracji)
SZW	– szyb windy
ŚC	– ściana
TN	– tolerowanie przez nadzór odstępstw od zasad bhp
UK	– utrata kontroli nad narzędziem
UKM	– utrata kontroli nad maszyną
UNO	– uderzenie o nieruchome czynniki
UPZ	– upadek do zagłębienia
UR	– urobek
URU	– uderzenie bądź pochwylenie przez elementy przemieszczane lub będące w ruchu
USP	– uderzenie przez spadające czynniki
UW	– upadek z wysokości
WY	– wybuch
WZ	– wejście, wjazd na obszar zagrożony
ZAS	– zasypanie w wykopie
ZN	– nagły przypadek
ZO	– niewłaściwa organizacja pracy

i ujednolicenie, celem zaimportowania do programu Statistica 12, wspomaganego środowiskiem R (2016) przy obliczaniu testów dokładnych. Arkusz „Wypadki” (tabela 3a–3c) zaimportowano do programu Statistica, gdzie w puste komórki (tabela 3c), oznaczające brak wystąpienia danego warunku w analizowanym wypadku, wpisano 0.

„Wypadki”. Źródło: opracowanie własne.

Tabela 3a

Lp	Sw	Wi (lata)	Stp (lata)	Ldn (dni)	Sz	Zwd	Mw
1	L	39	8	16	UPO	OPER. KOP.	BUDOWA
2	C	40	13	16	DG	PRAC. OGÓL.	BUDOWA

Tabela 3b

Pp	Cw	Cmc	Wdo	Cmo	Wyu	Cmu
RZ	PNZ	RM	POR	RM	POR	EI
TYNK	PNN	RU	UW	RU	UNO	PP

Tabela 3c

PRZYCZYNA														
BK	BŚI	BŚZ	NCM	NPD	POŚ	PNO	PUZ	SPF	SZ	TN	UKM	WZ	ZN	ZO
						1	1			1		1		1
	1	1	1					1	1					1

2.3. BADANIE WYPADKOWOŚCI NA TERENIE BUDOWY ZA POMOCĄ MODELI KLASYFIKACYJNYCH

W niniejszych badaniach wykorzystano dwie metody analiz zasobów danych: modelowanie klasyfikacyjne, w postaci drzew decyzyjnych (3.1) oraz grupowanie pojęciowe, w postaci analizy skupień (4.1). Podjęto w nich próbę wskazania zmiennych i wyznaczenia istotności ich udziału w zaistnieniu wypadków przy realizacji robót budowlanych.

2.3.1. DRZEWA KLASYFIKACYJNE

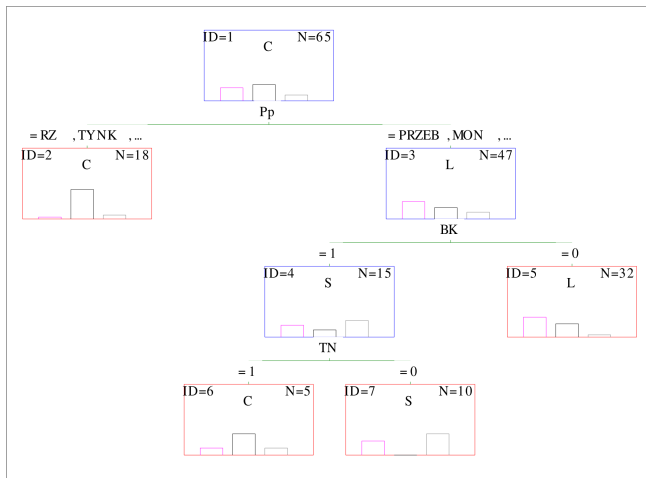
Drzewa klasyfikacyjne (decyzyjne) są jednym z rodzajów klasyfikatorów statystycznych. Wykorzystuje się do wyznaczenia przynależności przypadków lub obiektów do klas jakościowej zmiennej zależnej, na podstawie pomiarów jednej lub więcej zmiennych objaśniających (predyktorów). Inaczej mówiąc, drzewa decyzyjne starają się podzielić dane na jak najbardziej jednorodne grupy względem wartości zależnej. Należy zaznaczyć, że przy drzewach decyzyjnych nie występuje formalne testowanie statystycznej istotności zmiennych. Drzewa klasyfikacyjne dają się prosto przedstawiać graficznie, co sprawia, że są łatwiejsze w interpretacji, niż czysto liczbowe wyniki.

Analizę przeprowadzono dla arkusza „Wypadki” (65 obserwacji), ze zmienną zależną: *Sw* (skutek wypadku: L – lekki, C – ciężki, S – śmiertelny) oraz predyktorami (zmiennymi objaśniającymi): *Mw* (miejsce wypadku), *Pp* (proces pracy), *Cw* (czynność wykonywana przez poszkodowanego w chwili wypadku), *Cmc* (czynnik materialny związany z czynnością wykonywaną przez poszkodo-

wanego w chwili wypadku), *Wdo* (wydarzenie będące odchyleniem od stanu normalnego), *Cmo* (czynniki materialny związany z odchyleniem), *Wyu* (wydarzenie powodujące uraz), *Cmu* (czynniki materialny będący źródłem urazu).

Przy tworzeniu drzewa wykorzystano powszechnie stosowany algorytm CART (Classification and Regression Trees) – ciąg pytań, na które odpowiedzi determinują kolejne pytania, bądź kończą etap. W wyniku otrzymujemy strukturę drzewa, która w węzłach końcowych nie zawiera już pytań, lecz same odpowiedzi (Breiman i in., 1984).

Litera pojawiająca się w węźle (rys. 1) oznacza, jaki wypadek (L – lekki, C – ciężki i S – śmiertelny) został w tym węźle wyznaczony, np. w węźle 6 (ID = 6) wybrany został wypadek ze skutkiem ciężkim.



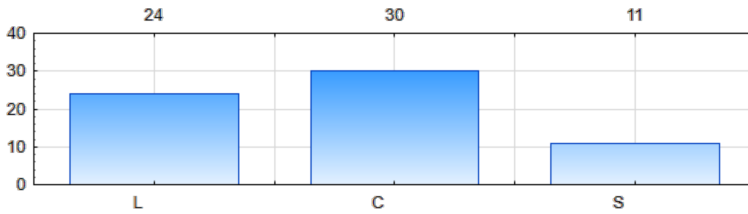
Rys. 1. Drzewo dla zmiennej zależnej Sw. Źródło: opracowanie własne.

Tabela 4.

Liczba wypadków zaklasyfikowanych do poszczególnych kategorii skutków wypadków, w węzłach końcowych drzewa dla zmiennej zależnej Sw – „Wypadki”. Źródło: opracowanie własne.

ID – węzeł	Klasa – L	Klasa – C	Klasa – S	Łącznie
2	1	15	2	18
6	1	3	1	5
7	4	0	6	10
5	18	12	2	32

Węzeł1 (ID = 1). Wśród wszystkich wypadków (tabela 4), najczęściej występowały wypadki skutkujące ciężkimi obrażeniami, najrzadziej zaś wypadki śmiertelne. Dokładny rozkład przedstawiono na rys. 2.



Rys. 2. Rozkład wypadków w węźle 1. Źródło: opracowanie własne.

Węzły 2 (ID = 2) i 3 (ID = 3), powstałe po podziale węzła 1 (ID = 1) ze względu na proces pracy (Pp). Podział ten przedstawia się następująco (tabela 5).

Tabela 5.

Podział węzła nr 1 i rozkład wypadków. Źródło: opracowanie własne.

Węzeł 2	Węzeł 3																
RZ (ziemne), TYNK, MUR	PRZEB, MON, DESK, SCH (przemieszczanie się), DEM, DACH, INST, TRANSP, BET																
<table border="1"> <thead> <tr> <th>Kategoria</th> <th>Liczba wypadków</th> </tr> </thead> <tbody> <tr> <td>L</td> <td>1</td> </tr> <tr> <td>C</td> <td>15</td> </tr> <tr> <td>S</td> <td>2</td> </tr> </tbody> </table>	Kategoria	Liczba wypadków	L	1	C	15	S	2	<table border="1"> <thead> <tr> <th>Kategoria</th> <th>Liczba wypadków</th> </tr> </thead> <tbody> <tr> <td>L</td> <td>23</td> </tr> <tr> <td>C</td> <td>15</td> </tr> <tr> <td>S</td> <td>9</td> </tr> </tbody> </table>	Kategoria	Liczba wypadków	L	23	C	15	S	9
Kategoria	Liczba wypadków																
L	1																
C	15																
S	2																
Kategoria	Liczba wypadków																
L	23																
C	15																
S	9																

Z tabeli 5 możemy odczytać, że przy robotach ziemnych, tynkarskich i mурowych, najczęściej występowały wypadki ciężkie (15 przypadków). W przypadku robót z węzła trzeciego (PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET), najwięcej było wypadków lekkich (23 przypadki), najmniej śmiertelnych (9 przypadków).

Węzły 4 (ID = 4) i 5 (ID = 5), powstały w wyniku podziału węzła 3 (ID = 3) na dwie gałęzie ze względu na posiadanie odpowiednich kwalifikacji (zmienna BK). Podział ten przedstawia się następująco (tabela 6).

Tabela 6.

Podział węzła nr 3 i rozkład wypadków. Źródło: opracowanie własne.

Węzeł 4	Węzeł 5																
Brak kwalifikacji (BK = 1)	Odpowiednie kwalifikacje (BK = 0)																
<table border="1"> <thead> <tr> <th>Kategoria</th> <th>Liczba wypadków</th> </tr> </thead> <tbody> <tr> <td>L</td> <td>5</td> </tr> <tr> <td>C</td> <td>3</td> </tr> <tr> <td>S</td> <td>7</td> </tr> </tbody> </table>	Kategoria	Liczba wypadków	L	5	C	3	S	7	<table border="1"> <thead> <tr> <th>Kategoria</th> <th>Liczba wypadków</th> </tr> </thead> <tbody> <tr> <td>L</td> <td>18</td> </tr> <tr> <td>C</td> <td>12</td> </tr> <tr> <td>S</td> <td>2</td> </tr> </tbody> </table>	Kategoria	Liczba wypadków	L	18	C	12	S	2
Kategoria	Liczba wypadków																
L	5																
C	3																
S	7																
Kategoria	Liczba wypadków																
L	18																
C	12																
S	2																

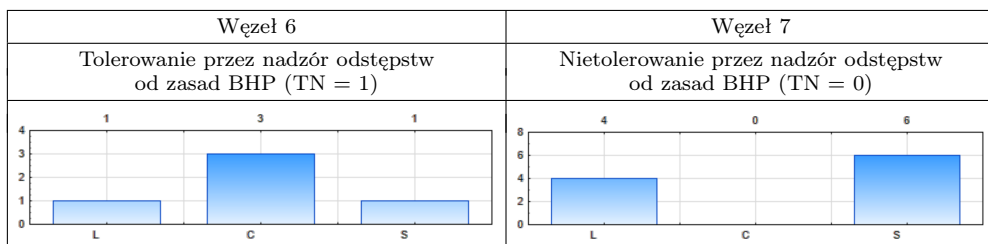
Z tabeli 6 wynika, że osoby wykonujące czynności zakwalifikowane do węzła 3 (PRZEB, MON, DESK, SCH (przemieszczanie się), DEM, DACH, INST,

TRANSP, BET), w przypadku braku kwalifikacji, najczęściej ulegają wypadkom ze skutkiem śmiertelnym. Z kolei wśród osób z odpowiednimi kwalifikacjami najczęściej pojawiały się wypadki ze skutkiem lekkim (18 przypadków), najrzadziej ze skutkiem śmiertelnym (2 przypadki).

Węzły 6 (ID = 6) i 7 (ID = 7), powstały w wyniku podziału węzła 4 (ID = 4) na dwie gałęzie ze względu na zmienną TN (tolerowanie przez nadzór odstępstw od zasad BHP). Podział ten przedstawia się następująco (tabela 7).

Tabela 7.

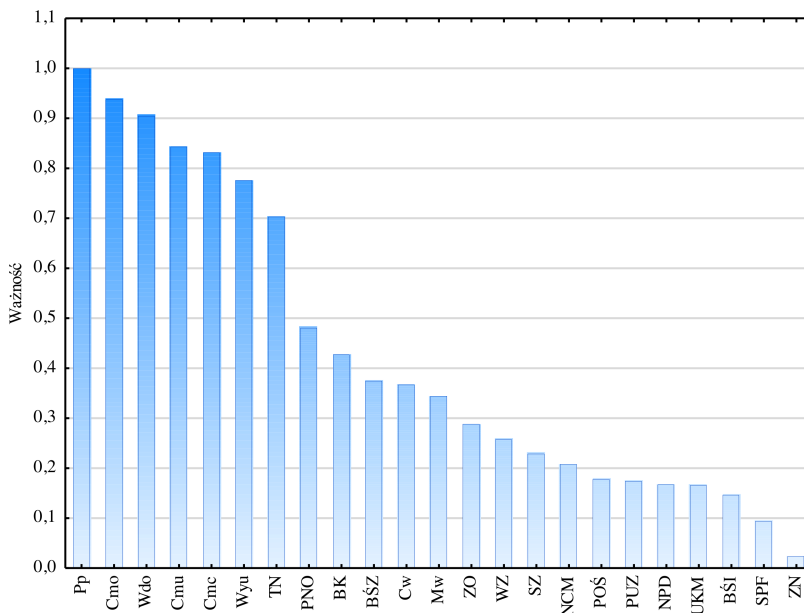
Podział węzła nr 4 i rozkład wypadków. Źródło: opracowanie własne.



Z tabeli 7 wynika, że brak kwalifikacji i tolerowanie przez nadzór odstępstw od zasad BHP najczęściej skutkuje wypadkami o skutku ciężkim. Z kolei w sytuacji, gdy nadzór nie toleruje odstępstw od zasad BHP, a pracownicy nie posiadają odpowiednich kwalifikacji, najczęściej występują wypadki śmiertelne (6 przypadków) i urazy lekkie (sporadyczne).

Dla analizowanego przypadku, najważniejszą zmienną różnicującą, jest proces pracy (Pp). Uwidacznia to rys. 3. Jednak istnieje jeszcze inne rozumienie ważności predyktorów. W STATISTICA zmienna może być najważniejszą, ale niekoniecznie musi być użyta w podziale (tzn. nie ma ani jednego podziału, w którym zmienna ta bierze udział). Wiąże się to ze sposobem obliczania ważności. Ważność predyktora obliczana jest na podstawie tego, jak jego uwzględnienie zwiększa „czystość” (ang. purity) otrzymanego podziału (podział jest najbardziej „czysty”, gdy w każdym węźle końcowym drzewa występuje tylko jedna wartość zmiennej zależnej, czyli udało się wyznaczyć reguły zawsze spełnione w zbiorze danych). Dla łatwiejszej interpretacji wykres ważności jest przeskalowany tak, aby najsilniejszy predyktor otrzymał wartość 1 i stanowił punkt odniesienia dla pozostałych.

Na przedstawionym wykresie (rys. 3) zaznacza się wyraźnie grupa najsilniej powiązanych ze skutkiem wypadku zmiennych – Cmo (czynnik materialny związany z odchyleniem), Wdo (wydarzenie będące odchyleniem od stanu normalnego), Cmu (czynnik materialny będący źródłem urazu), Cmc (czynnik materialny związany z czynnością wykonywaną przez poszkodowanego w chwili wypadku) oraz Wyu (wydarzenie powodujące uraz).



Rys. 3. Ważność predyktorów zmiennej zależnej Sw (L, C, S). Źródło: opracowanie własne.

Do oceny poprawności klasyfikacji zmiennych służy macierz klasyfikacji (tabela 8). Można ją utworzyć dla dowolnego rodzaju modeli klasyfikacyjnych. Dla przypadków występujących w zbiorze danych porównuje ona klasy wskazywane przez model z rzeczywistością zaobserwowanymi. Na tej podstawie można uzyskać nie tylko informację o ogólnej trafności przewidywań, ale też o najczęściej popełnianych rodzajach błędów.

Poniżej, na histogramie (rys. 4) przedstawiono wartości obserwowane i przewidywane dla danego modelu.

Z tabeli 8 wynika, że łącznie pojawiły się 24 wypadki ze skutkiem lekkim, z czego model poprawnie sklasyfikował 18, co stanowi 75%. Wypadki ciężkie wystąpiły 30 razy, z czego poprawnie sklasyfikowanych zostało 18, co stanowi 60%. Z kolei z 11 wypadków śmiertelnych, poprawnie rozpoznanych zostało 6, co stanowi 54,55%. Poprawność klasyfikacji zmiennych objaśniających przez drzewo (rys. 1) dla zmiennej niezależnej Sw (skutek wypadku) potwierdza histogram (rys. 4).

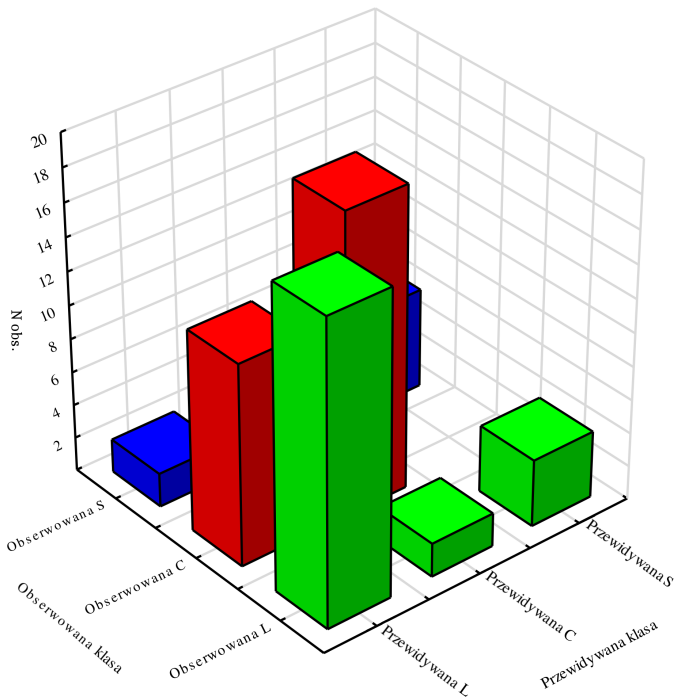
Reguły wynikające z powyższego drzewa przedstawiają się następująco:

- Jeżeli zmienna *Pp* (proces pracy) przyjęła wartość ze zbioru {RZ, TYNK, MUR} to najczęstszy skutek wypadku był ciężki (ID 2).
- Jeżeli zmienna *Pp* przyjęła wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik posiadał odpowiednie kwalifikacje, to najczęstszy skutek wypadku był lekki (ID 5).

Tabela 8.

Macierz klasyfikacji. Źródło: opracowanie własne.

	Przewidywana			Łącznie w wierszu
	L	C	S	
Obserwowana liczba wypadków L	18	2	4	24
Procent z kolumny	56,25%	8,70%	40,00%	
Procent z wiersza	75,00%	8,33%	16,67%	
Procent z ogółu	27,69%	3,08%	6,15%	36,92%
Obserwowana liczba wypadków C	12	18		30
Procent z kolumny	37,50%	78,26%	0,00%	
Procent z wiersza	40,00%	60,00%	0,00%	
Procent z ogółu	18,46%	27,69%	0,00%	46,15%
Obserwowana liczba wypadków S	2	3	6	11
Procent z kolumny	6,25%	13,04%	60,00%	
Procent z wiersza	18,18%	27,27%	54,55%	
Procent z ogółu	3,08%	4,62%	9,23%	16,92%
Obserwowana liczba ogółu wypadków	32	23	10	65
Procent łącznie	49,23%	35,38%	15,38%	



Rys. 4. Macierz klasyfikacji. Wartości obserwowane i przewidywane. Źródło: opracowanie własne.

- Jeżeli zmienna Pp przyjęła wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik nie posiadał odpowiednich kwalifikacji, a zmienna TN (tolerowanie przez nadzór odstępstw od zasad bhp) przyjęła wartość 1, to najczęstszy skutek wypadku był ciężki (ID 6).
- Jeżeli zmienna Pp przyjęła wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik nie posiadał odpowiednich kwalifikacji, a zmienna TN (tolerowanie przez nadzór odstępstw od zasad bhp) przyjęła wartość 0, to najczęstszy skutek wypadku był śmiertelny (ID 7).

2.4. BADANIE WYPADKOWOŚCI NA TERENIE BUDOWY ZA POMOCĄ GRUPOWANIA POJĘCIOWEGO

2.4.1. ANALIZA SKUPIEŃ

Analiza skupień jest narzędziem do eksploracyjnej analizy danych, której celem jest ułożenie elementów w grupy w taki sposób, aby stopień powiązania elementów z elementami należącymi do tej samej grupy był jak największy, a z elementami z pozostałych grup jak najmniejszy. Analiza skupień może być wykorzystywana do wykrywania struktur w danych bez wyprowadzania interpretacji, czy wyjaśnienia. Oznacza to, że analiza skupień jedynie wykrywa struktury w danych, bez wyjaśniania dlaczego one występują. Skupienia wyznaczone są w oparciu o pewne kryteria statystyczne, natomiast interpretacja wyników i ocena z praktycznego punktu widzenia należą do badacza (Breiman i in., 1984).

W niniejszym podrozdziale podjęto próbę wyznaczenia skupień (grup) zmiennych charakteryzujących wypadki budowlane zaistniałe w badanym miejscu i czasie. W efekcie końcowym analiza skupień pozwoli wyłonić zmienne, które odgrywały dominującą rolę w zaistnieniu wypadków budowlanych. Ich znajomość może być wykorzystana przy podejmowaniu decyzji w kontekście bezpiecznej realizacji robót budowlanych.

Ważnym problemem w tego rodzaju analizach jest określenie liczby poszukiwanych skupień. W programie Statistica możliwe jest określenie z góry tego parametru, przed uruchomieniem obliczeń, a następnie eksploracyjne dochodzenie do podziału o „ciekawej”, eksperckiej interpretacji. Można też wykorzystać automatyczną strategię poszukiwania podziału, z użyciem tzw. *v-krotnej* walidacji krzyżowej. Idea tej metody polega na podziale całej próby na v (w praktyce zwykle 5 lub 10) podzbiorów bądź losowych podprób rozłącznych. Ta sama analiza jest następnie stosowana do obserwacji z $v-1$ zbiorów (prób uczących), a wynikowy model stosowany do podzbioru v (próba lub podzbiór, który nie był wykorzystywany do wyznaczania skupień, tzn. próba testowa) i obliczana jest „trafność prognostyczna” (tzn. jak dobrze obserwacje z próby v są przypisywane

do jednorodnych skupień, za pomocą bieżącego rozwiązania obliczonego z $v-1$ prób uczących). Wyniki dla kolejnych v powtórzeń są agregowane (uśredniane) i dają jedną ocenę stabilności modelu, tzn. trafności przypisywania obserwacji do skupień.

Automatyczny wybór optymalnej liczby skupień jest realizowany w ten sposób, że tworzone są kolejno podziały na 2, 3 ... skupienia. Dla każdego z nich wyliczane jest – za pomocą walidacji krzyżowej – kryterium jakości podziału, po czym wybierany jest podział dający najlepszą wartość.

W niniejszych badaniach analizy przeprowadzono z wykorzystaniem modułu „Uogólniona analiza skupień”, programu Statistica, dla danych zawartych w arkuszu kalkulacyjnym „Wypadki”. Użyty modułoferuje dwa popularne algorytmy wyznaczania skupień (grup elementów tworzących jednorodne klasy):

- Algorytm k-średnich (k-means) – poszukuje takiego podziału, dla którego suma odpowiednio zdefiniowanych odległości pomiędzy skupieniami będzie jak największa w stosunku do średniej odległości w obrębie skupień (pomiędzy ich elementami). Algorytm ten można stosować do danych opisanych atrybutami numerycznymi.
- Algorytm EM (Expectation Maximisation; Witten, Frank 2000) zakłada, że dane pochodzą ze „zmieszania” próbek z rozkładów prawdopodobieństwa o różnych parametrach i stara się odtworzyć te rozkłady tak, aby uzyskać jak najlepsze dopasowanie. Można stosować go do danych numerycznych, ale też do zmiennych z kategoriami. W niniejszej analizie wykorzystano ten algorytm.

Zgromadzone do analiz dane o wypadkach są zarówno ilościowe, jak i jakościowe (skategoryzowane). W celu uzyskania większej spójności wszystkich wyników, zastosowano algorytm EM. Na bazie statystyk chi-kwadrat zbadano cechy najsilniej różnicujące otrzymane skupienia. Dla przyjętej segmentacji w arkuszach *Testy niezależności dla zmiennych jakościowych* (tabela 9) przedstawiono wyniki testu chi-kwadrat, oceniającego dla każdej zmiennej skategoryzowanej, czy jej rozkład jest istotnie różny dla różnych skupień. Wartość statystyki chi-kwadrat oblicza się według wzoru:

$$\chi^2 = \sum \frac{(O - E)^2}{E}, \quad (1)$$

gdzie:

E – oczekiwana częstość zmiennej,

O – obserwowana częstość zmiennej.

Wyniki istotne statystycznie zostały zaznaczone kursywą. Ponadto wielkość statystyki χ^2 oraz prawdopodobieństwo testowe możemy traktować jako miary tego, jak bardzo dana zmienna różnicuje skupienia – tym bardziej, im większa jest statystyka testowa/im mniejsze p .

W przypadku zastosowania zmiennych: Mw (miejsce wypadku), Pp (proces pracy), Cw (czynność wykonywana przez poszkodowanego), Cmc (czynnik materialny związany z Cw), Wdo (wydarzenie będące odchyleniem od stanu normalnego), Cmo (czynnik materialny związany z odchyleniem), Wyu (wydarzenie powodujące uraz), Cmu (czynnik materialny związany z Wdo) z arkusza „Wypadki” oraz dodatkowo zmienne metryczkowe z tego samego arkusza: Sz (status zatrudnienia), Wi (wiek poszkodowanego), Stp (staż pracy), Zwđ (zawód wykonywany) i algorytmu EM, Przeprowadzone analizy wykazały, że optymalna liczba skupień jest równa 2.

Tabela 9.

Testy niezależności dla zmiennych jakościowych „Wypadki”. Liczba skupień: 2. Całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienne	Chi-kwadrat ¹	Wartość p ²
Sz (status zatrudnienia)	3,64	0,6020
Zwđ (zawód wykonywany przez poszkodowanego)	31,86	0,0008
Mw (miejsce wypadku)	1,76	0,1842
Pp (proces pracy)	26,66	0,0052
Cw (czynność wykonywana przez poszkodowanego)	15,84	0,0032
Cmc (czynnik materialny związany z Cw)	59,04	0,0000
Wdo (wydarzenie będące odchyleniem od stanu normalnego)	41,58	0,0003
Cmo (czynnik materialny związany z Cw)	54,75	0,0000
Wyu (wydarzenie powodujące uraz)	38,50	0,0000
Cmu (czynnik materialny będący źródłem urazu)	41,42	0,0000
BK (brak wymaganych kwalifikacji)	0,14	0,7105
BŚI (brak środków ochrony indywidualnej)	1,79	0,1804
BŚZ (brak środków ochrony zbiorowej)	3,97	0,0465
NCM (niewłaściwy stan czynnika materialnego)	2,54	0,1113
NPD (nieodpowiednie przejścia i dojścia)	8,02	0,0046
POŚ (potknięcie i poślizg)	6,11	0,0134
PNO (wykonywanie prac z polecenia, nie z zakresu obowiązków)	0,56	0,4555
PUZ (wykonywanie prac bez usunięcia zagrożeń)	5,02	0,0251
SPF (niewłaściwy stan psycho – fizyczny)	0,24	0,6219
SZ (samowolne zachowanie pracownika)	2,14	0,1435
TN (tolerowanie przez nadzór odstępstw od zasad bhp)	1,05	0,3044
UKM (utrata kontroli nad maszyną)	3,87	0,0492
WZ (wejście, wjazd na obszar zagrożony)	5,17	0,0230
ZN (nagły przypadek)	9,23	0,0024
ZO (niewłaściwa organizacja pracy)	1,15	0,2843

¹Statystyka Chi-kwadrat Pearsona jest podstawą testu istotności zależności dla zmiennych jakościowych (skategoryzowanych). Miara ta oparta jest na możliwości obliczenia licznosci „oczekiwanych” w tabeli dwudzielczej (to znaczy licznosci, jakich oczekivalibyśmy, gdyby nie istniała zależność między zmiennymi).

²p-wartość – prawdopodobieństwo testowe (ang. p-value, probability value), graniczny poziom istotności (0,05) – miara prawdopodobieństwa popełnienia błędu pierwszego rodzaju, czyli liczbowe wyrażenie istotności statystycznej.

Im większa statystyka Chi-kwadrat, a mniejsza wartość p, tym bardziej skupienia 1 i 2 różnią się względem danej zmiennej.

Dla zmiennych ilościowych porównano średnie wartości oraz zakres w jednym i drugim skupieniu (tabela 10). Istotność różnic oceniono testem F analizy wariancji (ANOVA), w którym wartości statystyki F oblicza się według wzoru:

$$F = \text{stosunek} : \sum_{i=1}^K (\bar{Y}_i - \bar{Y})^2 / (K - 1) \quad \text{do} : \sum_{i=1}^K \cdot \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 / (N - K), \quad (2)$$

gdzie:

- n_i – liczba obserwacji w i -tej grupie,
- \bar{Y}_i – średnia wartość zmiennej w i -tej grupie,
- \bar{Y} – ogólna wartość zmiennych,
- Y_{ij} – obserwacja j w i z grupy K ,
- K – liczba grup,
- N – ogólna liczebność próby.

Tabela 10.

Wyniki testu ANOVA dla zmiennych ilościowych stosowanych w analizie „Wypadki”. Liczba skupień: 2. Całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienne	Międzygrupowe SS	df	Wewnątrzgrupowe SS	df	F	Wartość p
Wi (lata)	669,747	1	8478,71	63	4,98	0,0293
Stp (lata)	97,210	1	1709,93	63	3,58	0,0630

Wyznaczone skupienia różnią się istotnie zmienną: wiek poszkodowanych ($p = 0,0293$). Natomiast nie stwierdzono statystycznie różnic względem zmiennej staż pracy poszkodowanych (tabela 11).

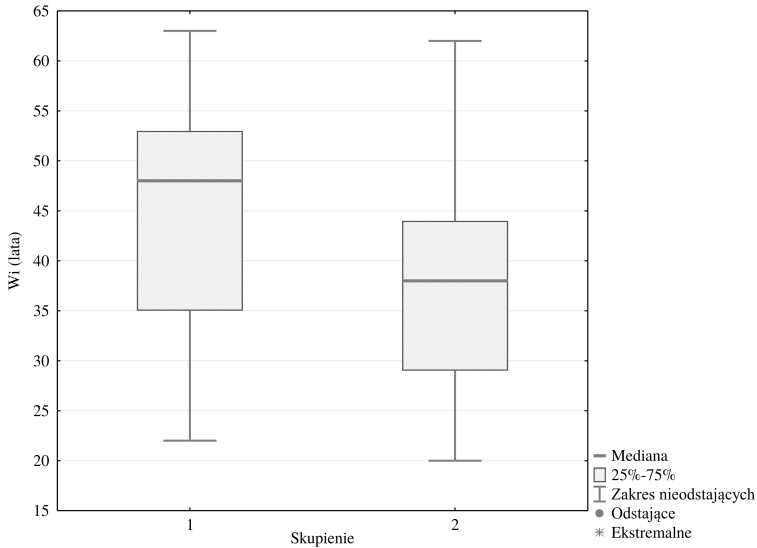
Tabela 11.

Statystyki zmiennej ilościowej: Wi (wiek poszkodowanych, w latach) – „wypadki 2014–16”. Liczba skupień: 2, całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

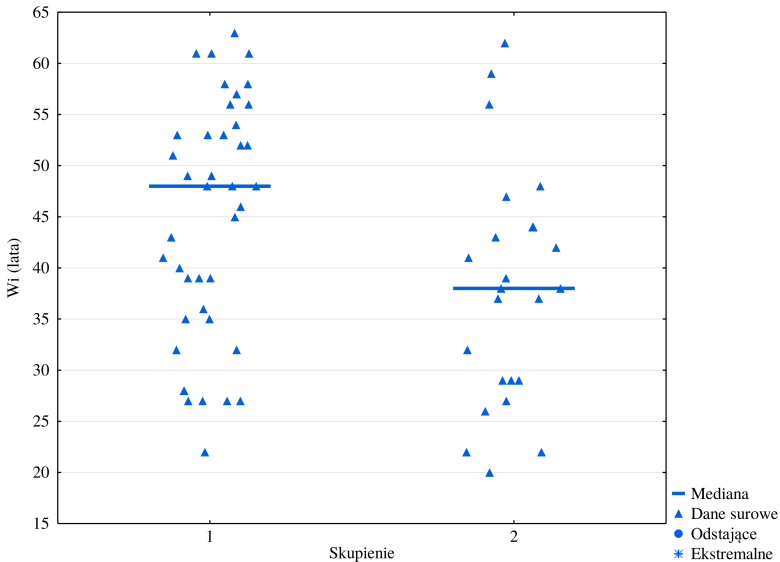
Wartości zmiennej Wi	Skupienie 1	Skupienie 2	Ogółem
Minimum	22,0	20,0	20,0
Maksimum	63,0	62,0	63,0
Średnia	44,6	38,0	42,2
Odchylenie std.	11,7	11,5	12,0

Najważniejsza dla interpretacji jest kolumna Wartość p (tabela 10), która wskazuje, że osoby ulegające wypadkom z pierwszego skupienia istotnie różnią się wiekiem od osób ze skupienia 2 (z tabeli 11 – statystyk opisowych możemy odczytać, że średni wiek w obu skupieniach to odpowiednio 44,6 i 38 lat), natomiast różnice w stażu pracy są mniej znaczące. Pozostałe wyniki w tabeli mają charakter raczej pomocniczy.

Wyniki zilustrowano też za pomocą wykresów ramka-wąsy oraz danych surowych. Pierwszy z nich (rys. 5) przedstawia graficznie przeciętną wartość, zakres zmienności i ewentualne wartości nietypowe badanej zmiennej, w obu skupieniach. Drugi (rys. 6) ilustruje, jak układają się surowe wartości zmiennej.



Rys. 5. Wykres ramkowy zmiennej ilościowej: Wi (wiek poszkodowanych) względem skupień 1 i 2. Klasyfikacja wynikowa. Źródło: opracowanie własne.



Rys. 6. Wykres danych surowych zmiennej ilościowej: Wi (wiek poszkodowanych) względem skupień 1 i 2. Źródło: opracowanie własne.

Poniżej, w tabelach: 12–15 i na rysunkach: 7–10, przedstawiono rozkłady poszczególnych zmiennych jakościowych w obrębie skupień: 1 i 2. Warunkiem włączenia zmiennej do wykresu było jej występowanie łącznie w obu skupieniach powyżej 5 krotnie.

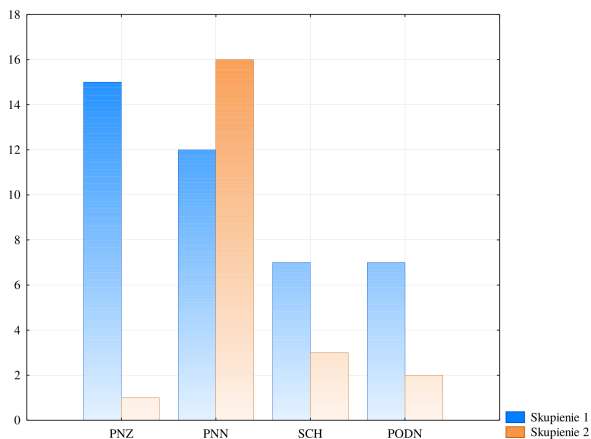
Rozkład zmiennej **Cw** – **czynność wykonywana przez poszkodowanego**, obejmował następujące zmienne (tabela 12).

Tabela 12.

Tabela licznosci dla zmiennej jakościowej Cw. Liczba skupień: 2, całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienna: Cw (czynność wykonywana przez poszkodowanego)			
Zmienne kodujące	Skupienie 1	Skupienie 2	Razem
PNZ (praca narzędziem zmechanizowanym)	15	1	16
PNN (praca narzędziem niezmechanizowanym)	12	16	28
SCH (przemieszczanie się (schodzenie, wchodzenie, przechodzenie))	7	3	10
PODN (transport ręczny (podnoszenie, opuszczanie))	7	2	9
BRAK	0	2	2

W rozkładzie zmiennej: **Cw** – **czynność wykonywana przez poszkodowanego**, w obrębie wyznaczonych skupień, dominują: **posługiwanie się narzędziem zmechanizowanym i niezmechanizowanym, przemieszczanie się i transport ręczny** (rys. 7).



Rys. 7. Wykres rozkładu zmiennej: Cw (czynność wykonywana przez poszkodowanego).

Źródło: opracowanie własne.

Rozkład zmiennej: **Cmc** – **czynnik materialny związany z czynnością wykonywaną**, obejmował następujące zmienne (tabela 13).

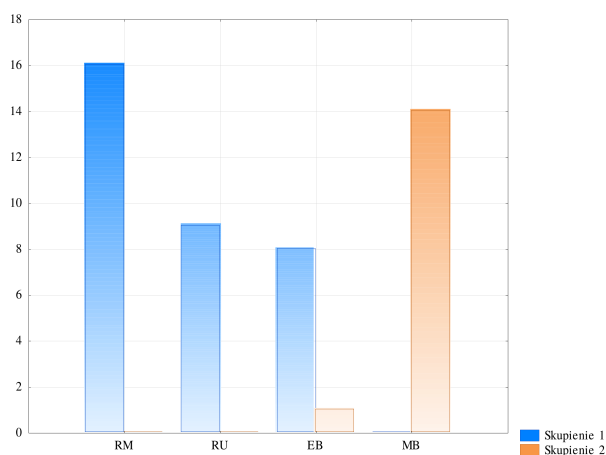
W rozkładzie zmiennej: **Cmc** – **czynnik materialny związany z czynnością wykonywaną**, w obrębie wyznaczonych skupień, dominują: **ruchome**

Tabela 13.

Tabela liczności dla zmiennej jakościowej Cmc. Liczba skupień: 2, całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienna: Cmc (czynnik materialny związany z czynnością wykonywaną)			
Zmienne kodujące	Skupienie 1	Skupienie 2	Razem
RM (ruchome maszyny i urządzenia)	16	0	16
RU (rusztowanie)	9	0	9
DP (drabina przenośna)	4	0	4
EB (element budynku)	8	1	9
MB (materiały budowlane)	0	14	14
SZW (szyb windowy)	0	2	2
EN (elektronarzędzie)	3	0	3
BRAK	0	2	2
EIg (element instalacji gazowej)	0	4	4
ŁO (łopata)	1	1	2

maszyny i urządzenia, rusztowania, element budynku i materiały budowlane (rys. 8).



Rys. 8. Wykres rozkładu zmiennej: Cmc (czynnik materialny związany z czynnością wykonywaną przez poszkodowanego). Źródło: opracowanie własne.

Rozkład zmiennej: **Wdo** – wydarzenie będące odchyleniem od stanu normalnego, obejmował następujące zmienne (tabela 14).

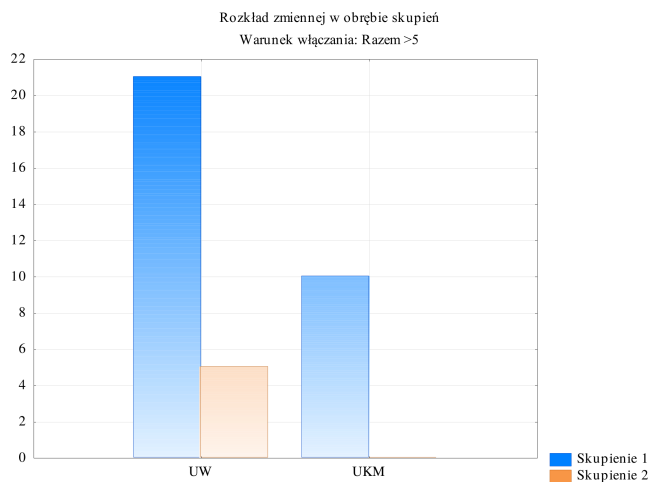
W rozkładzie zmiennej: **Wdo** – wydarzenie będące odchyleniem od stanu normalnego, w obrębie wyznaczonych skupień, dominują: **upadek z wysokości i utrata kontroli nad maszyną** (rys. 9).

Rozkład zmiennej: **Wyu** – wydarzenie powodujące uraz, obejmował następujące zmienne (tabela 15).

Tabela 14.

Tabela licznosci dla zmiennej jakościowej Wdo. Liczba skupień: 2, całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienna: Wdo (wydarzenie będące odchyleniem od stanu normalnego)			
Zmienne kodujące	Skupienie 1	Skupienie 2	Razem
POR (porażenie prądem)	1	0	1
UW (upadek z wysokości)	21	5	26
UK (utrata kontroli nad narzędziem)	1	3	4
UKM (utrata kontroli nad maszyną)	10	0	10
POŚ (potknięcie i poślizg)	1	0	1
NIEOKR (nieokreślony)	0	1	1
WY (wybuch)	0	2	2
PRZ (przygnięcie)	0	3	3
SP (spadający przedmiot)	0	1	1
ZN (nagły przypadek)	0	2	2
KN (kontakt z narzędziem)	0	2	2
UPZ (upadek do zagłębienia)	5	0	5
UNO (uderzenie o nieruchome czynniki)	0	2	2
USP (uderzenie przez spadające czynniki)	0	1	1
OE (przecięcie i skaleczenie)	1	2	3
ZAS (zasypanie w wykopie)	1	0	1



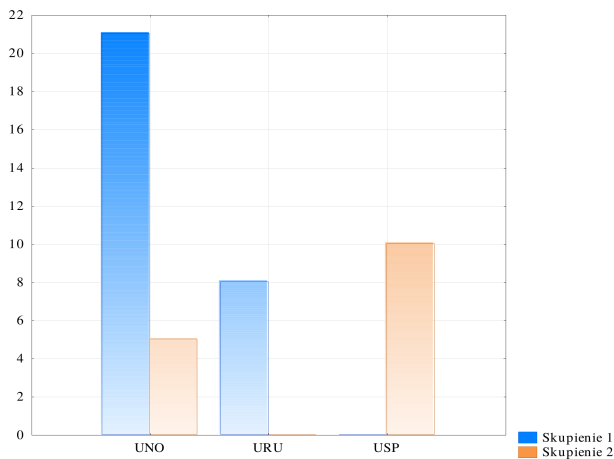
Rys. 9. Wykres rozkładu zmiennej: Wdo (wydarzenie będące odchyleniem od stanu normalnego). Źródło: opracowanie własne.

W rozkładzie zmiennej: **Wyu** – **wydarzenie powodujące uraz**, w obrębie wyznaczonych skupień, dominują: **uderzenie o nieruchome czynniki, uderzenie bądź pochwycenie przez elementy przemieszczane lub będące w ruchu i uderzenie przez spadające czynniki** (rys. 10).

Tabela 15.

Tabela licznosci dla zmiennej jakościowej Wyu. Liczba skupień: 2, całkowita liczba przypadków uczących: 65. Źródło: opracowanie własne.

Zmienna: Wyu (wydarzenie powodujące uraz)			
Zmienne kodujące	Skupienie 1	Skupienie 2	Razem
POR (porażenie prądem)	1	2	3
UNO (uderzenie o nieruchome czynniki)	21	5	26
URU (uderzenie bądź pochwylenie przez elementy przemieszczane lub będące w ruchu)	8	0	8
OMS (przeciążenie układu mięśniowo-szkieletowego)	4	1	5
KN (kontakt z narzędziem)	0	1	1
WY (wybuch)	0	2	2
USP (uderzenie przez spadające czynniki)	0	10	10
UPZ (upadek do zagłębienia)	5	0	5
OE (przecięcie i skaleczenie)	1	2	3
PRZ (przygniecenie)	0	1	1
ZAS (zasypanie w wykopie)	1	0	1



Rys. 10. Wykres rozkładu zmiennej: Wyu (wydarzenie powodujące uraz).

Źródło: opracowanie własne.

Przeprowadzona analiza skupień pozwoliła wyznaczyć grupy zmiennych (skupienia) charakteryzujących wypadki budowlane, zaistniałe w badanym miejscu i czasie. Spośród wyznaczonych grup wyłoniono zmienne, które odgrywały dominującą rolę w wyznaczonych skupieniach.

Wykorzystując dane z rejestrów powypadkowych, prowadzonych przez Okręgowy Inspektorat Pracy (arkusz „Wypadki”): Mw (miejsce wypadku), Pp (proces pracy), Cw (czynność wykonywana), Cmc (czynnik materialny związany z czynnością wykonywaną), Wdo (wydarzenie będące odchyleniem od stanu normalne-

go), Cmo (czynnik materialny związany z odchyleniem), Wyu (wydarzenie powodujące uraz), Cmu (czynnik materialny będący źródłem urazu) oraz dodatkowo zmienne metryczkowe, z tego samego arkusza: Sz (status zatrudnienia), Wi (wiek poszkodowanego), Stp (staż pracy poszkodowanego), Zwd (zawód wykonywany) i algorytm EM, przeprowadzone analizy umożliwiły wyróżnienie dwóch skupień (grup) wartości zmiennych opisujących najważniejsze przyczyny wypadków. Różnią one istotnie wiekiem osób poszkodowanych (średni wiek w obu skupieniach to 44,6 i 38 lat), natomiast różnice w stażu pracy i wartościach pozostałych zmiennych objaśniających są mniej znaczące.

2.5. PODSUMOWANIE

Wypadki przy pracy, jako zdarzenia na terenie budowy, są zjawiskami losowymi, trudnymi lub niemożliwymi do przewidzenia. Dlatego ich badanie i wskazanie związków, między charakteryzującymi je cechami, nie jest łatwe. Dotyczy to w szczególności analiz danych szczegółowych, w których przedmiotem badań są czynniki generujące niepożądane następstwa dla życia i zdrowia ludzi, w związku z podejmowaną pracą. W niniejszym rozdziale podjęto próbę naukowej analizy wielowymiarowego zbioru danych, opisanego przez cechy ilościowe i jakościowe, które charakteryzują wypadki przy pracy podczas realizacji robót budowlanych. Przedmiotem analiz były tereny budów, na których miały miejsce wypadki przy pracy, odnotowane w rejestrze Okręgowego Inspektoratu Pracy w Krakowie, w latach 2014 i 2016.

Zasadniczym celem niniejszego opracowania była eksploracja wpływu wybranych czynników na bezpieczeństwo realizacji robót budowlanych, wyrażonego przez zachowanie pracownika, rodzaj i status warunków do zaistnienia wypadku oraz rozpatrzenie przydatności nowoczesnych technik ekstrakcji danych do realizacji tej analizy.

W zakresie metodologicznym posłużono się dwoma sposobami analizowania zasobów danych. Pierwszy sposób należy do kategorii modeli nadzorowanych (supervised) – objaśniających wskazaną zmienną zależną (drzewa decyzyjne). Drugi sposób należy do kategorii modeli nienadzorowanych (unsupervised) – nastawionych na grupowanie powiązanych ze sobą danych różnego typu w jednym obszarze (analiza skupień).

Drzewa decyzyjne są dobrym narzędziem do wyznaczania przynależności przypadków lub obiektów do klas jakościowej zmiennej zależnej, na podstawie pomiarów jednej lub więcej zmiennych objaśniających (predyktorów). Starają się podzielić dane na jak najbardziej jednorodne grupy względem wartości zmiennej zależnej. Nie przeprowadza się przy nich testowania statystycznej istotności zmiennych. Dają prosty, graficzny, obraz, co sprawia, że są łatwiejsze w interpretacji, niż czysto liczbowe wyniki. Na podstawie przeprowadzonych analiz wykryto następujące reguły (zależności przyczynowo-skutkowe):

- Jeżeli zmienna Pp (proces pracy) przyjmuje wartość ze zbioru {RZ, TYNK, MUR} to najczęstszy skutek wypadku jest ciężki.
- Jeżeli zmienna Pp przyjmuje wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik posiada odpowiednie kwalifikacje, to najczęstszy skutek wypadku jest lekki.
- Jeżeli zmienna Pp przyjmuje wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik nie posiada odpowiednich kwalifikacji, a zmienna TN (tolerowanie przez nadzór odstępstw od zasad bhp) przyjmuje wartość 1, to najczęstszy skutek wypadku jest ciężki.
- Jeżeli zmienna Pp przyjmuje wartość ze zbioru {PRZEB, MON, DESK, SCH, DEM, DACH, INST, TRANSP, BET} i pracownik nie posiada odpowiednich kwalifikacji, a zmienna TN (tolerowanie przez nadzór odstępstw od zasad bhp) przyjmuje wartość 0, to najczęstszy skutek wypadku jest śmiertelny.

Również analiza skupień, mimo narzuconych ograniczeń, daje obszerny zbiór wyników. Za pomocą algorytmu EM wyznaczono podział zbioru wypadków na dwa podzbiory. Analizując rozkład wykorzystanych w podziale zmiennych, w obrębie utworzonych skupień, wyróżniono ich następujące charakterystyki:

Skupienie 1

- Przyczynami wypadków są: utrata kontroli nad maszyną, praca bez usunięcia zagrożeń, wjazd na obszar zagrożony.
- Wypadki przy pracy narzędziami zmechanizowanymi.
- Typowe czynniki mechaniczne „występujące” w wypadku: ruchome maszyny, rusztowania, elementy budynku.

Skupienie 2

- Przyczynami wypadków są: nagłe przypadki, potknięcia, brak środków ochrony zbiorowej, nieodpowiednie przejścia i dojścia.
- Wypadkom ulegają osoby poniżej 40 roku życia.
- Głównie wypadki przy pracy narzędziami niezmechanizowanymi.
- Wypadki, w których czynnikiem są materiały budowlane.

Wykorzystując zmienne z arkusza „Wypadki”: Mw (miejsce wypadku), Pp (proces pracy), Cw (czynność wykonywana przez poszkodowanego), Cmc (czynnik materialny związany z czynnością wykonywaną przez poszkodowanego w chwili wypadku), Wdo (wydarzenie będące odchyleniem od stanu normalnego), Cmo (czynnik materialny związany z odchyleniem), Wyu (wydarzenie powodujące uraz), Cmu (czynnik materialny będący źródłem urazu) oraz dodatkowo zmienne metryczkowe, z tego samego arkusza: Sz (status zatrudnienia), Wi (wiek poszkodowanego), Stp (staż pracy poszkodowanego), Zwd (zawód wykonywany przez poszkodowanego), przeprowadzone analizy umożliwiły wyróżnienie dwóch skupień (grup) wartości zmiennych opisujących najważniejsze przyczyny wypadków.

Różnią one istotnie wiekiem osób poszkodowanych (średni wiek w obu skupieniach to 44,6 i 38 lat), natomiast różnice w stażu pracy i wartościach pozostałych zmiennych objaśniających są mniej znaczące.

Zaproponowane w niniejszym rozdziale podejście, sposób analizy i weryfikacji pozwalają na zauważenie i opisanie czynników wpływających na wypadkowość w budownictwie do tej pory nie badanych albo badanych inaczej. Otrzymane wyniki mogą być wykorzystane przez projektantów i kierowników budów na etapie sporządzania informacji dotyczącej bezpieczeństwa i ochrony zdrowia oraz planu bezpieczeństwa i ochrony zdrowia, jak również na etapie projektowania organizacji robót i późniejszego kierowania ich realizacją. W szczególności wyniki te mogą być pomocne w procesie decyzyjnym, przy opracowywaniu różnego rodzaju działań, mających na celu poprawę bezpieczeństwa pracy w budownictwie i redukcję liczby wypadków.

2.6. BIBLIOGRAFIA

- BREIMAN L., FRIEDMAN J., STONE C.J., OLSHEN R.A. (1984). *Classification and Regression Trees*. Wadsworth Statistics/Probability.
- CHOUDHRY R.M., FANG D. (2008). Why operatives engage in unsafe work behavior: Investigating factors on construction sites, *Safety Science*, 46, 566–584.
- DROZD W. (2016). Charakterystyka terenu budowy w aspekcie zagrożeń bezpieczeństwa pracy, *Czasopismo Inżynierii Lądowej, Środowiska i Architektury*, Politechnika Rzeszowska, styczeń – marzec, 165–172.
- DROZD W. (2017). *Badania cech terenu budowy i ich wpływu na bezpieczeństwo prowadzenia robót budowlanych przy obiektach nieliniowych*, Monografia, Politechnika Krakowska, Kraków.
- HOŁA B., SZOSTAK M. (2015). Analysis of the state of the accident rate in the construction industry in European Union Countries, *Archives of Civil Engineering*, 4, 19–34.
- HOŁA B. (2016). *Bezpieczeństwo pracy w procesach budowlanych*, Wydawnictwo: Politechnika Wroclawska, Wrocław.
- SHAPIRA A., ASCE F., LYACHIN B. (2009). Identification and Analysis of Factors Affecting Safety on Construction Sites with Tower Cranes, *Journal of Construction Engineering and Management*, 135.
- TÖRNER M., POUSETTE P. (2009). Safety in construction – a comprehensive description of the characteristics of high safety standards in construction work, from the combined perspective of supervisors and experienced workers, *Journal of Safety Research*, 40, 399–409.