

MARIA MARKS<sup>1)</sup>

## WYKORZYSTANIE METOD UCZENIA MASZYNOWEGO DO ROZPOZNANIA ZWIĄZKÓW MIĘDZY SKŁADEM BETONU MODYFIKOWANEGO DODATKAMI POPIOŁOWYMI A JEGO WŁAŚCIWOŚCIAMI

**STRESZCZENIE.** W pracy wykorzystano dwa algorytmy uczenia maszynowego do określenia związków między składem betonu modyfikowanego popiołami ze spalania węgla kamiennego i brunatnego, a trwałością na agresywne oddziaływanie środowiska. Do określenia przenikalności chlorków w betonie użyto przyspieszonej metody migracji ujętej w Nordtest Method NT Build 492. Do określenia odporności betonu na powierzchniowe łuszczenie zastosowano zgodną z normą szwedzką metodę Borås. W obu przypadkach zgromadzone dane doświadczalne wykorzystano jako zbiory uczące, na podstawie których wygenerowano reguły. Reguły wygenerowane za pomocą algorytmu AQ21 oraz algorytmu J48 z systemu WEKA pozwalają na zakwalifikowanie betonów zwykłych oraz betonów, w których część cementu została zastąpiona popiołami fluidalnymi z węgla kamiennego lub brunatnego, do grupy betonów o dobrej lub dostatecznej odporności na wnikanie chlorków oraz odpornych i nieodpornych na powierzchniowe łuszczenie, spowodowane cyklicznym zamrażaniem i odmrażaniem.

### 1. WPROWADZENIE

Z powodu dużej różnorodności składników współczesnych kompozytów cementowych, stosowanych jednocześnie i w różnych ilościach, związki między tymi składnikami

---

<sup>1)</sup> dr hab. – Instytut Podstawowych Problemów Techniki PAN, Warszawa

oraz trwałością materiałów wykraczają poza możliwości prostego opisu inżynierskiego. Znane i tworzone modele materiałów ujmują wspomniane związki na ogół w formie empirycznej, ograniczonej do przetestowanego zestawu materiałów, przy intuicyjnym wyborze czynników dominujących. Na przykład: analizując wpływ dodatków popiołowych na wytrzymałość czy trwałość betonu pomija się na ogół wpływ zmiennej konsystencji mieszanki lub zmiennej zawartości domieszek uplastyczniających, przyjmując, że jest pomijalnie mały. Zagadnienie komplikuje się w przypadku wielu parametrów składu (różnych właściwości cementu i jego zawartości, różnych właściwości i zawartości domieszek chemicznych, włókien, ziaren kruszywa), warunków wykonania i pielęgnacji, a także agresywnych oddziaływań środowiskowych. Sformułowanie modelu empirycznego w takiej wielowymiarowej przestrzeni parametrów składu i właściwości kompozytu jest zbyt skomplikowane. W takiej sytuacji do obiektywnego rozpoznania związków między składem a właściwościami kompozytu można wykorzystać metody komputerowe, zwane metodami sztucznej inteligencji.

Stosowanie metod sztucznej inteligencji, takich jak metody sztucznych sieci neuronowych czy metody uczenia maszynowego wymagają przygotowania odpowiednich baz danych. Bazę danych stanowi wyselekcjonowany i uporządkowany zestaw informacji na temat materiału, technologii, eksploatacji tego materiału, jego cech i właściwości. Bazę danych można przedstawić jako macierz złożoną z wierszy i kolumn, które wypełniają nie tylko liczby, ale także charakterystyki jakościowe, nazwy, zmienne boolowskie itp. Dane mogą być uzyskiwane na podstawie opracowań i raportów z laboratoriów polskich oraz zagranicznych, jak również z artykułów w czasopiśmie. Istotne jest odpowiednie określenie struktury baz danych oraz analiza danych eksperymentalnych. Wyjściowym problemem jest uzgodnienie jednolitej formy rejestracji i przekazywania danych.

Dane o kompozycie składają się z danych wejściowych, stanowiących informację o tym co wiadomo o danym materiale i jego wykonaniu, (input), oraz dane wyjściowe, (output), dotyczące cech, właściwości, które mogą być przedmiotem analizy, predykcji, itd. W poszczególnych zadaniach rozpoznawania wiedzy te same wielkości mogą występować w odmiennej roli. Na przykład 28-dniowa wytrzymałość na ściskanie betonu może być zmienną rozpoznawaną (ang. *output*), przy predykcji na podstawie wiedzy o składzie mieszanki, lub też może być zmienną wejściową (ang. *input*), np. przy przewidywaniu trwałości betonu.

Wspomniane sztuczne sieci neuronowe (SSN, SN, ANNS – ang. *Artificial Neural Networks*), opracowane jako programy komputerowe, są układami służącymi do przekształcania danych liczbowych w sposób, który jest wzorowany na funkcjonowaniu komórek nerwowych w mózgu. Sztuczne sieci neuronowe realizują rozmaite algorytmy uczenia, przy których stany pamięci układu, tak zwane wagi, ulegają adaptacji odpowiednio do napływających informacji. Dostarczane informacje muszą wstępnie zostać przekształcone do postaci numerycznej (na przykład przez przypisanie kolejnych liczb rozmaitych stosowanym rodzajom cementu). Cechą wyróżniającą sztuczne sieci neuronowe jest specyficzne odtwarzanie zależności pomiędzy informacją wejściową a informacją wyjściową.

Sieci SSN różnią się w istotny sposób od konwencjonalnych programów komputerowych, które można określić jako sekwencyjne. W programie sekwencyjnym należy przewidzieć wszystkie ewentualności na jakie program ma reagować. Sieci neuronowe stanowią wygodną i tanią propozycję wieloprocesorowego systemu o bardzo wielu elementach przetwarzających równolegle dostarczane informacje. W sieciach neuronowych ostateczny sposób działania programu ustala się dopiero w wyniku procesu nauczania. Sieć zapamiętuje dostarczane informacje w sposób podobny do układu biologicznego. Każdy neuron dysponuje pewną wewnętrzną pamięcią, reprezentowaną przez aktualne wartości wag, oraz pewnymi możliwościami przetwarzania wejściowych sygnałów w sygnał wyjściowy. Z tak zdefiniowanych neuronów konstruowana jest sieć neuronowa. Struktura sieci powstaje w ten sposób, że wyjścia jednych neuronów łączy się (według wybranego schematu) z wejściami innych tworząc łącznie system zdolny do równoległego przetwarzania różnych informacji. Charakterystyka sieci neuronowej zależy w znacznym stopniu od sposobu połączenia neuronów, czyli tzw. architektury sieci.

W działaniu sieci neuronowej można wyodrębnić etap nauki, kiedy sieć gromadzi informacje potrzebne jej do określenia, co i jak ma robić, oraz etap normalnego działania podczas, którego opierając się o zdobytą wiedzę sieć rozwiązuje konkretne zadania.

Ilustracją wykorzystania sztucznych sieci neuronowych do predykcji kompozytów betonopodobnych modyfikowanych pyłami krzemionkowymi jest przedstawione w pracy [1] prognozowanie wytrzymałości 28 dniowego betonu na ściskanie. Predykcję tą dokonano za pomocą trzech następujących programów sztucznych sieci neuronowych:

- programu **Beton** uzyskanego za pomocą implementacji sieci FuzzyARTMAP [2],
- programu **ssn6** opartego na koncepcji wstecznej propagacji błędu czyli sieci typu BP [3 - 4],
- programu **aiNet**, który bazuje na systemie samoorganizującym się (ang. self organizing system) i jest zbliżony do metody najbliższych sąsiadów [5].

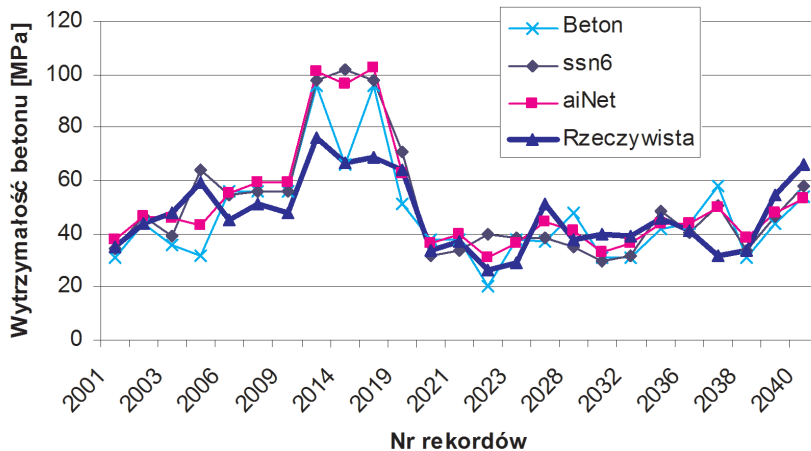
Analizowana baza danych, stanowiąca zbiór uczący, składa się z 788 rekordów zawierających dane o wartościach 6 atrybutów wejściowych:

- $C$  – zawartość cementu [ $\text{kg}/\text{m}^3$ ],
- $W$  – zawartość wody [ $\text{kg}/\text{m}^3$ ],
- $Su$  – zawartość superplastyfikatora [ $\text{kg}/\text{m}^3$ ],
- $SF$  – zawartość pyłów krzemionkowych [ $\text{kg}/\text{m}^3$ ],
- $FA$  – zawartość kruszywa drobnego [ $\text{kg}/\text{m}^3$ ],
- $CA$  – zawartość kruszywa grubego [ $\text{kg}/\text{m}^3$ ],

oraz jednego atrybutu wyjściowego:

- $fc28$  – wytrzymałość 28 dniowa betonu na ściskanie [MPa].

Przyjętą predykcję wytrzymałości betonu przeprowadzono na zbiorze danych składającym się z 25 rekordów, które dotyczyły reprezentatywnych mieszanek wybranych spośród 40 wykonanych w IPPT PAN w latach 2000 - 2004 i oznaczonych na osi odciętych numerami od 2001 do 2040. Na rysunku 1 zestawiono wyniki predykcji wytrzymałości przy zastosowaniu programów **Beton**, **ssn6** i **aiNet** oraz wartości pomierzone na kostkach  $150 \times 150 \times 150$  mm.



Rys. 1. Zestawienie wyników predykcji wytrzymałości uzyskanych przy zastosowaniu trzech różnych sieci neuronowych. Linia pogrubioną zaznaczono wytrzymałość rzeczywistą  
Fig. 1. Predicted values of the strength, obtained by three types of neural networks.  
The thick line denotes the measured strength

Dokonując zestawienia wyników predykcji zastosowanych programów i wytrzymałości rzeczywistej obliczono wielkości średnich błędów, które w programie **Beton** wynoszą 21%, w programie **ssn6** 20% i w programie **aiNet** 18%.

Sieci neuronowe przetwarzają liczby, w których starają się wykryć struktury stanowiące podstawę do późniejszego rozpoznawania innych podobnych danych. Ponadto, sieci neuronowe funkcjonują na danych kompletnych, w których występują wszystkie wartości atrybutów, a dane są liczbami rzeczywistymi. W przypadku, gdy wartości niektórych atrybutów nie są określone, to muszą być uzupełnione albo też trzeba po prostu z tych rekordów zrezygnować.

## 2. METODY UCZENIA MASZYNOWEGO

Algorytmy z zakresu uczenia maszynowego podobnie jak sieci neuronowe należą do kategorii układów uczących się. Celem algorytmów uczenia maszynowego jest eksploracja danych tzn. wydobywanie informacji ukrytej w dostarczonym lub zgromadzonym zbiorze danych. Zbiór danych jest przedstawiany jako tzw. baza danych, tzn. zbiór rekordów nazwanych także przykładami lub przypadkami opisanych za pomocą

ustalonego zestawu atrybutów. Odkrywanie regularności w danych, polegające na generowaniu, na podstawie analizy dostarczonych danych trenujących, hipotezy, która stanowi ogólny opis występujących w tych danych zależności jest uczeniem indukcyjnym, ściśle opartym na przykładach [6]. Wygenerowana hipoteza może z kolei być stosowana do wnioskowania dedukcyjnego w odniesieniu do nowych danych celem ich weryfikacji.

Jednym z najczęściej rozwiązywanych problemów metodami uczenia maszynowego jest zagadnienie klasyfikacji. W takim przypadku wybrany atrybut reprezentuje klasę, a metody odkrywania wiedzy służą poszukiwaniu zależności pozwalających na określenie przynależności przykładów do konkretnej klasy na podstawie pozostałych atrybutów np. określenie odporności betonu na agresywne oddziaływanie jonów chlorkowych na podstawie atrybutów opisujących skład i właściwości betonu. W zależności od tego, w jaki sposób algorytm uczenia maszynowego odkrywa wspomniane zależności można wyróżnić algorytmy oparte o tworzenie drzew decyzyjnych (np. algorytm C4.5) bądź reguł decyzyjnych (rodzina algorytmów AQ). Niezależnie od sposobu działania metod odkrywania wiedzy, ich wspólną cechą jest tworzenie hipotez.

Wygenerowana na podstawie zbioru uczącego hipoteza, w postaci np. drzewa decyzyjnego lub zbioru reguł określana jako klasyfikator, powinna dobrze przybliżać klasyfikację spoza tego zbioru. Wygenerowany klasyfikator można ocenić za pomocą różnych kryteriów, między innymi trafności klasyfikowania, złożoności obliczeniowej algorytmów, wymogów co do minimalnej liczby przykładów niezbędnych do akceptowalnego przybliżenia klasyfikatora.

W przypadku klasyfikatorów ocena trafności klasyfikowania jest podstawowym narzędziem weryfikacji. Ocena trafności klasyfikowania przeprowadzana jest na zbiorze przykładów testowych, różnych od przykładów trenujących, za pomocą klasyfikatora wygenerowanego na podstawie danych uczących. Dla przykładów testowych znana jest ich rzeczywista przynależność do klas. Pozwala to na porównanie klasy uzyskanej w wyniku działania klasyfikatora z rzeczywistością i stwierdzenie, czy klasyfikacja danego przykładu jest poprawna czy błędna. Miarą trafności klasyfikowania jest dokładność klasyfikacji (ang. *classification accuracy*), która jest stosunkiem liczby poprawnie sklasyfikowanych przykładów testowych do wszystkich przykładów testowych. Miara ta wyrażona jest w procentach. Im większa wartość dokładności klasyfikacji tym skuteczniejszy jest klasyfikator. Na bardziej wszechstronną analizę pomyłek w przynależności do różnych klas przykładów testowych pozwala tzw. macierz pomyłek, inaczej niezgodności (ang. *confusion matrix*). Jest to macierz kwadratowa, w której wiersze odpowiadają rzeczywistym, a kolumny prognozowanym klasom, tzn. na przecięciu  $i$ -tego wiersza i  $j$ -tej kolumny występuje liczba przykładów testujących należących do rzeczywistej  $i$ -tej klasy a zakwalifikowanej przez klasyfikator do  $j$ -tej klasy. Na podstawie macierzy niezgodności można określić dokładność klasyfikacji jako stosunek sumy liczb na przekątnej (suma liczb przypisanych  $i$ -temu wierszowi i  $i$ -tej kolumnie przy  $i = 1, \dots, m$ , gdzie  $m$  liczba klas), które są przykładami poprawnie zaklasyfikowanymi do liczby wszystkich przykładów testowych.

W przypadku, gdy systemowi uczącemu dostarczony zostanie tylko jeden zbiór przykładów, należy podzielić ten zbiór w sposób losowy na część uczącą i testującą. Typowy podział to 2/3 przykładów – część ucząca, a 1/3 to część testująca. Jednakże najczęściej stosowaną techniką podziału jest  $k$ -krotna walidacja wskrośna, inaczej ocena krzyżowa (ang. *k-fold cross validation*), która dzieli losowo cały zbiór przykładów na  $k$  równolicznych podzbiorów  $U = E_1 \cup \dots \cup E_k$  [3]. W każdej  $i$ -tej iteracji zbiór  $E_i$  jest zbiorem testowym, a zbiór  $U \setminus E_i$  jest zbiorem uczącym. Trafność klasyfikowania jest wyliczana jako wartość średnia z dokładności klasyfikowania w każdej iteracji.

Aby zilustrować w jaki sposób algorytm uczenia maszynowego odkrywa zależności pozwalające określić przynależność do konkretnej klasy na podstawie dostarczonych atrybutów, dokonano określenia odporności betonu na agresywne oddziaływanie jonów chlorkowych oraz mrozu i soli odladzającej wykorzystując atrybuty opisujące skład i właściwości betonu. W tym celu zastosowano:

- program uczenia maszynowego AQ21 [7], który jest najnowszym przedstawicielem rodziny systemów AQ.

Celem wnioskowania w programie AQ21 jest wyznaczenie hipotez indukcyjnych wyrażonych w postaci jak najbardziej zrozumiałej dla człowieka, najbliższej językowi naturalnemu. Algorytm AQ, podobnie jak większość algorytmów uczenia się reguł, jest oparty na prostym schemacie sekwencyjnego pokrywania. A zatem polega on na kolejnym generowaniu reguł, aż do momentu pokrycia całego (czasami prawie całego) zbioru przykładów trenujących. W tym celu ze zbioru niepokrytych dotychczas przykładów wybierany jest jeden, dla którego poszukiwana jest reguła, która jednocześnie pokrywa jak najwięcej przykładów z tej samej klasy (tzw. przykłady pozytywne) nie pokrywająca przykładów z innych klas (tzw. przykłady negatywne),

- algorytm J48, udostępniony w systemie WEKA.

System Weka [8] jest zbiorem bardzo wielu algorytmów z zakresu uczenia maszyn. Został opracowany przez zespół programistów z Uniwersytetu Waikato w Nowej Zelandii. Wykorzystany algorytm J48 jest implementacją ostatniej publicznie dostępnej wersji algorytmu indukcji drzew decyzyjnych C4.5. Algorytm indukcji drzew decyzyjnych oparty jest na schemacie zstępującego konstruowania drzewa [3], [8]. Zasadniczym problemem jest wybór atrybutu do zbudowania testu, na podstawie którego nastąpi w węźle podział zbioru przykładów. Dobrym testem jest ten, którego użycie w węźle powoduje skrócenie ścieżki prowadzącej przez ten węzeł do liści wskazujących klasę decyzyjną. Z wygenerowanego przez algorytm J48 drzewa decyzyjnego dość łatwo można przejść do reguł.

### 3. GENEROWANIE REGUŁ OKREŚLAJĄCYCH ODPORNOŚĆ BETONU MODYFIKOWANEGO POPIOŁAMI FLUIDALNYMI NA PRZENIKANIE CHLORKÓW

#### 3.1. ODPORNOŚĆ BETONU NA PRZENIKANIE CHLORKÓW PO 28 DNIACH

W celu określenia reguł opisujących odporność betonu na przenikanie chlorków wykorzystano wybrane atrybuty dotyczące składu, zawartości powietrza w mieszance, wytrzymałości na ściskanie po 28 dniach oraz współczynnika migracji chlorków po 28 dniach dojrzewania betonów trzech serii, w których 15% i 30% cementu lub 20% i 40% w ostatniej serii było zastąpione popiołem fluidalnym.

Do określenia przenikalności chlorków w betonie użyto przyspieszoną metodę migracji ujętą w Nordtest Method NT Build 492 [9]. Współczynnik migracji chlorków w betonach określany był przy nieustalonym przepływie strumienia chlorków. Przyjęto następujące kryteria oceny współczynnika migracji chlorków w betonie zaproponowane przez L. Tanga [10] (tabl. 1).

Tablica 1. Ocena odporności betonu na wnikanie chlorków  
Table 1. Estimation of concrete resistance to chloride penetration

Współczynnik migracji	Odporność na wnikanie chlorków
$< 2 \cdot 10^{-12} \text{ m}^2/\text{s}$	bardzo dobra
$2 \cdot 10^{-12} \div 8 \cdot 10^{-12} \text{ m}^2/\text{s}$	dobra
$8 \cdot 10^{-12} \div 16 \cdot 10^{-12} \text{ m}^2/\text{s}$	dopuszczalna
$> 16 \cdot 10^{-12} \text{ m}^2/\text{s}$	niedopuszczalna

Na podstawie kryterium Tanga dokonano oceny odporności betonu na przenikanie chlorków po 28 dniach w zależności od wartości współczynnika migracji chlorków, [11]. W tablicy 2 przedstawiono zgromadzoną bazę danych składającą się z 15 rekordów, zawierającą 6 atrybutów numerycznych i jeden atrybut nominalny. Przyjęto następujące oznaczenia:

- $C1$  – zawartość cementu [ $\text{kg}/\text{m}^3$ ],
- $pfT$  – zawartość fluidalnego popiołu lotnego z węgla brunatnego (z Turowa) [ $\text{kg}/\text{m}^3$ ],
- $pfK$  – zawartość fluidalnego popiołu lotnego z węgla kamiennego (z Katowic) [ $\text{kg}/\text{m}^3$ ],
- $W$  – ilość wody [ $\text{kg}/\text{m}^3$ ],
- $A_{fr}$  – zawartość powietrza w mieszance betonowej [%],
- $fc28$  – wytrzymałość 28 dniowa betonu na ściskanie [MPa].

Ostatni atrybut jest atrybutem nominalnym przyjętym jako klasa dwuwartościowa (dobra, dopuszczalna). W rozpatrywanej bazie danych do klasy [odporność na przenikanie chlorków – dobra] należy 9 przykładów, a do klasy [odporność na przenikanie chlorków – dopuszczalna] należy 6 przykładów.

Tablica 2. Baza danych  
Table 2. The database

<i>C1</i>	<i>pfT</i>	<i>pfK</i>	<i>W</i>	<i>A_fr</i>	<i>fc28</i>	Odporność na przenikanie chlorków
360	0	0	162	2,1	55,0	dopuszczalna
306	0	54	162	1,8	56,2	dopuszczalna
252	0	108	162	1,3	51,6	dobra
306	54	0	162	1,6	60,3	dobra
252	108	0	162	1,6	58,7	dobra
380	0	0	171	6,2	46,3	dopuszczalna
323	0	57	171	6,8	47,2	dobra
266	0	114	171	5,8	46,8	dobra
323	57	0	171	6,6	45,3	dopuszczalna
266	114	0	171	6,2	46,3	dopuszczalna
406	0	0	175	4,9	22,7	dopuszczalna
290	73	0	151	6,9	21,0	dobra
217	145	0	150	7,8	26,1	dobra
323	0	81	167	4,6	38,3	dobra
244	0	162	157	4,6	43,0	dobra

Celem poszukiwań jest wygenerowanie reguł, które na podstawie dostępnych atrybutów pozwalałyby określić odporność betonu na przenikanie jonów chlorkowych. W tym celu wykorzystano algorytm J48. Jako zbiór uczący przyjęto wszystkie przykłady zgromadzone w bazie danych. Drzewo decyzyjne wygenerowane przez algorytm J48 przedstawione jest na rysunku 2, gdzie pierwsza liczba w nawiasie określa liczbę przykładów ze zbioru uczącego pokrytych przez daną regułę, a druga po znaku „,” wskazuje liczbę niepoprawnie zakwalifikowanych przykładów do tej klasy, czyli przykłady negatywne. Jeśli zatem występuje tylko jedna liczba to określa ona liczbę przykładów poprawnie zakwalifikowanych do tej klasy czyli przykłady pozytywne.

Na podstawie otrzymanego drzewa decyzyjnego łatwo można uzyskać reguły następującej postaci:

[Odporność na przenikanie chlorków – dobra]

reguła1 [ $C1 \leq 323$ ] i [ $pfK \leq 54$ ] i [ $W \leq 162$ ],

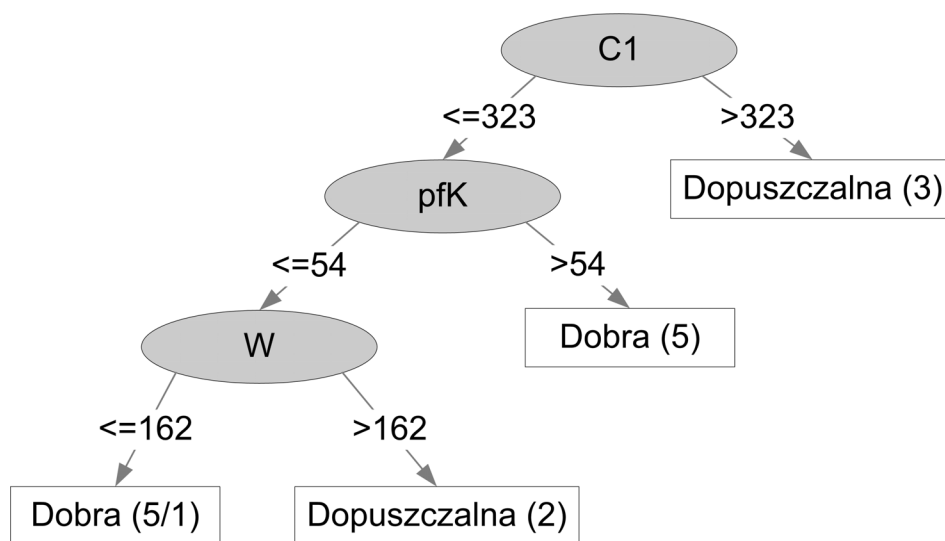
reguła2 [ $C1 \leq 323$ ] i [ $pfK > 54$ ],

[Odporność na przenikanie chlorków – dopuszczalna]

reguła1 [ $C1 \leq 323$ ] i [ $pfK \leq 54$ ] i [ $W > 162$ ],

reguła2 [ $C1 > 323$ ].





Rys. 2. Drzewo decyzyjne wygenerowane przez algorytm J48 określające odporność betonu na przenikanie chlorków po 28 dniach

Fig. 2. The decision tree generated by J48 algorithm describing the resistance of concrete to chloride penetration after 28 days of curing

Oszacowując wygenerowane reguły na zbiorze testowym pokrywającym się ze zbiorem uczącym otrzymano następującą macierz niezgodności:

	Dobra	Dopuszczalna
Dobra	9	0
Dopuszczalna	1	5

z której wynika, że jeden przykład z klasy dopuszczalnej jest błędnie klasyfikowany do klasy dobrej, pozostałe 14 przykładów jest klasyfikowanych poprawnie, czyli dokładność klasyfikacji wynosi 93,3%.

Właściwa ocena jakości klasyfikatora wymaga jednak, aby został on przetestowany na innym zbiorze danych niż dane uczące – tzw. zbiorze testowym. Jeśli nie dysponujemy zbiorem testowym wtedy najczęściej do oceny trafności klasyfikacji stosowana jest technika  $k$ -krotnej walidacji wskrośnej. Gdy baza danych składa się z małej liczby przykładów (mniej niż 100), wskazane jest stosowanie  $k$ -krotnej walidacji wskrośnej, w której liczba iteracji jest równa liczbie przykładów. Zatem przyjmując  $k = 15$  otrzymano dokładność klasyfikacji wynoszącą 60%, a macierz niezgodności przyjmuje następującą postać:

	Dobra	Dopuszczalna
Dobra	7	2
Dopuszczalna	4	2

### 3.2. ODPORNOŚĆ BETONU NA PRZENIKANIE CHLORKÓW PO 90 DNIACH

W celu określenia reguł opisujących odporność betonu na przenikanie chlorków po 90 dniach zastosowano program AQ21 i wykorzystano bazę danych analogiczną do bazy przedstawionej w tabelicy 2. Pierwsze 5 atrybutów numerycznych jest takich samych jak w tabelicy 1, ostatni atrybut numeryczny  $f_{c90}$  określa wytrzymałość betonu na ściskanie po 90 dniach. W zamieszczonej bazie danych do klasy [odporność na przenikanie chlorków – dobra] należy 12 przykładów, a do klasy [odporność na przenikanie chlorków – dopuszczalna] należą 3 przykłady (tabl. 3).

Tablica 3. Baza danych  
Table 3. The database

$C1$	$pfT$	$pfK$	$W$	$A_{fr}$	$f_{c90}$	Odporność na przenikanie chlorków
360	0	0	162	2,1	70,0	dopuszczalna
306	0	54	162	1,8	64,3	dobra
252	0	108	162	1,3	61,0	dobra
306	54	0	162	1,6	70,4	dobra
252	108	0	162	1,6	66,3	dobra
380	0	0	171	6,2	49,8	dopuszczalna
323	0	57	171	6,8	48,4	dobra
266	0	114	171	5,8	56,4	dobra
323	57	0	171	6,6	50,1	dobra
266	114	0	171	6,2	47,7	dobra
406	0	0	175	4,9	26,3	dopuszczalna
290	73	0	151	6,9	23,3	dobra
217	145	0	150	7,8	25,3	dobra
323	0	81	167	4,6	41,8	dobra
244	0	162	157	4,6	43,4	dobra

Jako zbiór uczący przyjęto wszystkie przykłady zgromadzone w bazie. Program AQ21 wygenerował następujące reguły:

[Odporność na przenikanie chlorków – dobra]

reguła1 [ $C1 \leq 341$ ]:  $p = 12$ ,  $n = 0$ ,  $q = 1$ ,

reguła2 [ $C1 \leq 351$ ] i [ $f_{c90} \leq 68,15$ ]:  $p = 11$ ,  $n = 0$ ,  $q = 0,987$ ,

[Odporność na przenikanie chlorków – dopuszczalna]

reguła1 [ $C1 \geq 342$ ]:  $p = 3$ ,  $n = 0$ ,  $q = 1$ ,

gdzie:  $p$  określa liczbę rekordów pozytywnych,  $n$  liczbę rekordów negatywnych tzn. liczbę rekordów spełniających tę regułę z klasy innej, a  $q$  określa jakość reguły.

Powyższe reguły można interpretować w następujący sposób pamiętając, że dotyczą one betonów o łącznej wartości materiałów wiążących wynoszącej 360, 380 i 406  $[\text{kg}/\text{m}^3]$  (tabl. 3):

[Odporność na przenikanie chlorków jest dobra]

JEŚLI

$[C1 \leq 341]$

LUB

$[C1 \leq 351]$  i  $[f_{c90} \leq 68,15]$

[Odporność jest dopuszczalna]

JEŚLI

$[C1 \geq 342]$ .

Wygenerowane reguły zweryfikowane na zbiorze przykładów uczących wykazują pełną poprawność klasyfikacji przykładów, co dobrze ilustruje macierz niezgodności:

	Dobra	Dopuszczalna
Dobra	9	0
Dopuszczalna	0	6

Do oceny trafności klasyfikacji zastosowano również metodę k-krotnej walidacji wskrośnej przy  $k = 15$ , której wyniki można przedstawić w postaci następującej macierzy niezgodności:

	Dobra	Dopuszczalna
Dobra	11	1
Dopuszczalna	0	3

Z macierzy niezgodności wynika, że jeden przykład z klasy dobrej został nieprawidłowo zakwalifikowany do klasy dopuszczalnej natomiast 14 przykładów jest prawidłowo zakwalifikowanych, a zatem dokładność klasyfikacji wynosi 93,3%.

Z wygenerowanych reguł określających odporność betonu na przenikanie chlorków po 90 dniach wynika, że do klasy dopuszczalnej należą betony, w których zawartość popiołu fluidalnego w spoiwie jest zerowa. Natomiast do klasy dobrej należą betony, w których część cementu od 15% do 40% została zastąpiona popiołem fluidalnym z węgla kamiennego lub z węgla brunatnego.

#### 4. GENEROWANIE REGUŁ OKREŚLAJĄCYCH ODPORNOŚĆ BETONU MODYFIKOWANEGO POPIOŁAMI FLUIDALNYMI NA POWIERZCHNIOWE ŁUSZCZENIE

Do określenia odporności betonu na powierzchniowe łuszczenie zgodnie z normą PKN- CEN/TS 12390-9:2007 należy wyznaczyć średnie wartości masy złuszczonego materiału z górnej powierzchni próbek po 28 cyklach  $m_{28}$  i po 56 cyklach  $m_{56}$  zamrażania i odmrażania w obecności 3% roztworu soli. Na podstawie tych wartości można dokonać oceny odporności betonu na powierzchniowe łuszczenie stosując szwedzką normę SS 13 72 44 [12] tzw. metodę Borås. Metoda Borås wyróżnia cztery kategorie betonów, w zależności od ilości złuszczonego materiału oraz od szybkości postępujących złuszczeń:

- beton bardzo dobrej jakości:  $m_{56} \leq 0,10 \text{ kg / m}^2$ ,
- beton dobrej jakości:  $m_{56} \leq 0,20 \text{ kg / m}^2$   
lub  $m_{56} < 0,50 \text{ kg / m}^2$  i  $m_{56}/m_{28} \leq 2$ ,
- beton dopuszczalnej jakości:  $m_{56} \leq 1,00 \text{ kg / m}^2$  i  $m_{56}/m_{28} < 2$ ,
- beton niedopuszczalnej jakości:  $m_{56} > 1,00 \text{ kg / m}^2$  i  $m_{56}/m_{28} > 2$ .

Ze względu na małą liczebnie bazę danych, która składa się z 17 rekordów, wprowadzono tylko dwie klasy odporności na łuszczenie powierzchniowe (TAK, NIE). Do klasy „TAK” przyjęto według klasyfikacji metody Borås beton bardzo dobrej, dobrej i dopuszczalnej jakości, natomiast do klasy „NIE” przyjęto beton niedopuszczalnej jakości. W zgromadzonej bazie danych wybrano atrybuty dotyczące czterech serii betonów napowietrzonych. W serii pierwszej 15% i 30% masy cementu zastąpiono popiołem fluidalnym z węgla brunatnego i kamiennego, w serii drugiej 20% i 40% cementu zastąpiono obydwoma popiołami, w serii trzeciej 20%, 30% i 40% cementu zastąpiono popiołem fluidalnym z węgla kamiennego a w serii ostatniej 30% cementu zastąpione popiołem z węgla brunatnego i kamiennego. Zgromadzona baza danych zawierająca 6 atrybutów numerycznych i jeden nominalny przedstawiona jest w tabeli 4. Przyjęto następujące oznaczenia:

- $C1$  – zawartość cementu [ $\text{kg/m}^3$ ],
- $pfT$  – zawartość fluidalnego popiołu lotnego z węgla brunatnego [ $\text{kg/m}^3$ ],
- $pfK$  – zawartość fluidalnego popiołu lotnego z węgla kamiennego [ $\text{kg/m}^3$ ],
- $w/s$  – stosunek wody do spoiwa,
- $A_{hr}$  – zawartość powietrza w betonie [%],
- $L$  – wskaźnik rozmieszczenia porów powietrznych [mm].

Tablica 4. Baza danych  
Table 4. The database

$C1$	$pfT$	$pfK$	$w/s$	$A_{hr}$	$L$	Odporność na łuszczenie powierzchniowe
380	0	0	0,45	4,46	0,38	TAK
323	0	57	0,45	4,83	0,28	NIE
266	0	114	0,45	4,33	0,35	NIE
323	57	0	0,45	4,70	0,34	NIE
266	114	0	0,45	6,88	0,24	NIE
406	0	0	0,42	10,08	0,07	TAK
323	0	81	0,42	5,94	0,12	TAK
290	73	0	0,42	18,41	0,05	TAK
244	0	162	0,42	6,07	0,14	NIE
217	145	0	0,42	16,56	0,08	TAK
360	0	0	0,42	6,25	0,13	TAK
288	0	72	0,42	6,25	0,13	NIE
252	0	108	0,42	7,08	0,14	NIE
216	0	144	0,42	6,03	0,18	NIE
354	0	0	0,44	3,70	0,17	TAK
246	0	105	0,44	2,00	0,27	NIE
246	105	0	0,44	3,80	0,56	NIE

Atrybuty  $A_{hr}$  i  $L$  są parametrami charakteryzującymi strukturę napowietrzenia betonu. Ostatni atrybut „Odporność na łuszczenie powierzchniowe” jest atrybutem nominalnym, przyjętym jako klasa dwuwartościowa (TAK, NIE). W rozpatrywanej bazie do klasy [Odporność na łuszczenie powierzchniowe – TAK] należy 7 przykładów, a do klasy [Odporność na łuszczenie powierzchniowe – NIE] należy 10 przykładów.

Do wygenerowania reguł pozwalających na podstawie atrybutów określić odporność betonu na łuszczenie powierzchniowe wykorzystano algorytm J48, a jako zbiór uczący przyjęto również wszystkie przykłady zgromadzone w bazie danych (tabl. 4). Drzewo decyzyjne wygenerowane przez algorytm J48 przedstawione jest na rysunku 3, gdzie liczba w nawiasie oznacza liczbę przykładów poprawnie zakwalifikowanych do tej klasy.

Na podstawie otrzymanego drzewa decyzyjnego łatwo można uzyskać reguły następującej postaci:

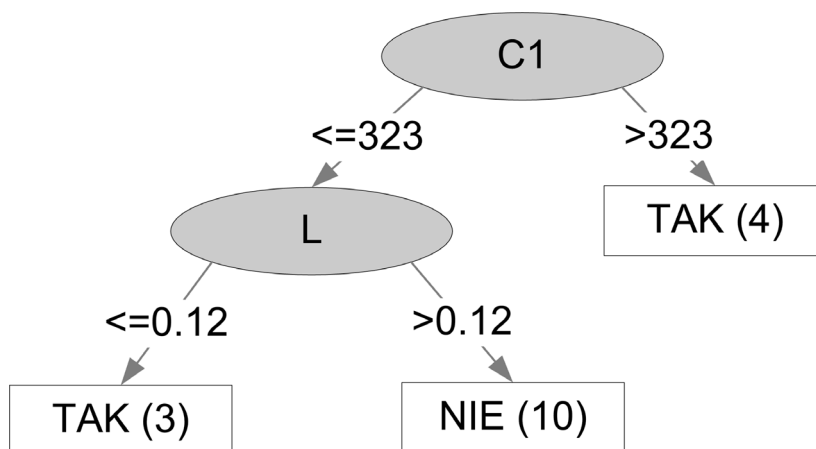
[Odporność na łuszczenie powierzchniowe – TAK]

reguła1 [ $C1 \leq 323$ ] i [ $L \leq 0,12$ ]:  $p = 3$ ,  $n = 0$ ,

reguła2 [ $C1 > 323$ ]:  $p = 4$ ,  $n = 0$ ,

[Odporność na łuszczenie powierzchniowe – NIE]

reguła1 [ $C1 \leq 323$ ] i [ $L > 0,12$ ]:  $p = 10$ ,  $n = 0$ .



Rys. 3. Drzewo decyzyjne wygenerowane przez algorytm J48 określające odporność betonu na łuszczenie powierzchniowe

Fig. 3. The decision tree generated by J48 algorithm describing the surface scaling resistance of concrete

Reguły weryfikowane na zbiorze przykładów uczących wykazują dokładność klasyfikacji wynoszącą 100%, co przedstawia następująca macierz niezgodności:

	TAK	NIE
TAK	7	0
NIE	0	10

Do oceny trafności klasyfikacji została zastosowana również metoda k-krotnej walidacji wskrośnej przy  $k = 17$  (wartość parametru  $k$  pokrywa się z liczbą przykładów), której wyniki można przedstawić w postaci następującej macierzy niezgodności:

	TAK	NIE
TAK	5	2
NIE	3	7

Dokładność klasyfikacji wynosi zatem 70,6%.

## 5. PODSUMOWANIE

W przedstawionym opracowaniu stosując dwa algorytmy uczenia maszynowego: algorytm AQ21 i algorytm J48 z Pakietu WEKA wygenerowano reguły określające:

- odporność betonów z dodatkiem popiołów fluidalnych na agresywne oddziaływanie chlorków po 28 i 90 dniach oraz
- odporność betonów z dodatkiem popiołów fluidalnych na powierzchniowe łuszczenie po 28 i 56 cyklach zamrażania i odmrażania w obecności soli.

Reguły wygenerowane za pomocą programu AQ21 oraz algorytmu J48 z Pakietu WEKA pozwalają na zakwalifikowanie betonów zwykłych oraz betonów, w których część cementu od 15% do 40% została zastąpiona popiołami fluidalnymi z węgla kamiennego lub brunatnego, do grupy betonów o dobrej lub dostatecznej odporności na wnikanie chlorków oraz do odpornych i nieodpornych na powierzchniowe łuszczenie. Na podstawie wygenerowanych reguł wynika, że najistotniejszymi atrybutami pozwalającymi określić odporność betonu na agresywne oddziaływanie chlorków okazały się zawartość cementu i zawartość popiołu fluidalnego, natomiast w przypadku odporności na powierzchniowe łuszczenie do tych dwóch wymienionych atrybutów dochodzi wskaźnik rozmieszczenia porów powietrznych, który jest parametrem charakteryzującym strukturę napowietrzenia betonu.

Ze względu na małą liczbę próbek reguły te mają jedynie zastosowanie do betonów o zbliżonej ilości materiałów wiążących oraz zbliżonym stosunku wodno - spoiwowym.

Niezbędne są dalsze badania w celu powiększenia bazy danych. Do uzyskania reguł pozwalających na ocenę odporności na wnikanie chlorków jak również odporności na powierzchniowe łuszczenie różnych betonów stosowanych w budownictwie niezbędne jest przeprowadzenie badań doświadczalnych na znacznie większej liczbie próbek wykonanych z betonów o różnej ilości materiałów wiążących, różnym składzie kruszywa i różnych stosunkach *w/s*. Wykorzystując tak rozbudowaną bazę można wygenerować reguły pozwalające na wstępną ocenę odporności na wnikanie chlorków jak również odporności na powierzchniowe łuszczenie, bez przeprowadzania czasochłonnych i kosztownych badań.

## BIBLIOGRAFIA

- [1] *Alterman D.*: Ocena materiałów betonowych przy zastosowaniu wnioskowania automatycznego. Rozprawa doktorska, IPPT PAN, Warszawa 2005
- [2] *Carpenter G.A., Grossberg S., Markuzon N., Reynolds J.H., Rosen D.B.*: Fuzzy ARTMAP: a natural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transaction of Neural Networks*, **3**, 5, September 1992, 698 - 713
- [3] *Krawiec K., Stefanowski J.*: *Uczenie maszynowe i sieci neuronowe*. Wydawnictwo Politechniki Poznańskiej, 2003
- [4] *Kasperkiewicz J., Alterman D.*: *Rozpoznawanie właściwości betonu metodami sztucznej inteligencji: Metody diagnozowania betonów i betonów wysokowartościowych na podstawie badań strukturalnych*. IPPT PAN, Warszawa 2003
- [5] aiNet: A neural network application for Windows environment. Version 1.24, User's Guide, Slovenia 1997, ([www.ainet-sp.si](http://www.ainet-sp.si)), 68
- [6] *Cichosz P.*: *Systemy uczące się*. Warszawa, WNT, 2000
- [7] *Wojtusiak J.*: AQ21 user's guide. George Mason University, MLI 04-3, September 2004

- [8] *Witten I.H., Frank E.:* Data mining. Practical machine learning tools and techniques, Elsevier, 2005
- [9] Nordtest Method NT Build 492. Concrete, mortar and cement-based repair materials: Chloride migration coefficient from non-steady-state migration experiments, 1999
- [10] *Tang L.:* Chloride transport in concrete – Measurement and prediction. Publication P-96:6, Chalmers University of Technology, Department of Building Materials, Göteborg, 1996
- [11] *Marks M., Józwiak-Niedźwiedzka D., Glinicki M.A.:* Application of machine learning for prediction of concrete resistance to migration of chlorides, Brittle Matrix Composites 9, Warszawa 25 - 28.10.2009, Woodhead Publishing Ltd & IPPT, 227 - 236
- [12] SS 137244:2005 Concrete testing - Hardened concrete - Scaling at freezing. Swedish Standard Institute, Stockholm 1995

## **INFORMACJE DODATKOWE**

Praca współfinansowana ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Operacyjnego Innowacyjna Gospodarka, nr projektu: POIG.01.01.02-24-005/09

## **APPLICATION OF MACHINE LEARNING TO DETERMINE PROPERTIES OF CONCRETES MODIFIED WITH ADDITION OF COAL ASH**

### **Abstract**

In the paper two algorithms of the machine learning are used in order to determine the durability of concrete modified with circulating fluidized bed combustion (CFBC) ash from hard coal and from brown coal. The rapid chloride permeability test, according to Nordtest Method BUILD 492, was used for determining the chloride ions penetration in concrete. The frost salt scaling tests were performed according to the Swedish Standard method SS 137244. In both cases the performed tests provided databases used as training sets to generate the rules describing the relations between material composition and durability parameters. The rules generated by computer programs AQ21 and WEKA using J48 algorithm provided means for adequate categorization of plain concrete and concrete modified with CFBC fly ash as materials of good and acceptable resistance to chloride penetration as well as materials resistant or not resistant to the surface scaling caused by freezing cycles.