



Predykcja poziomu tlenu w piecu EAF z wykorzystaniem metod sztucznej inteligencji: regresji liniowej i metody najbliższych sąsiadów (k-NN)

Using artificial intelligence for prediction of oxygen content in electric arc furnaces (EAF): linear regression and k-Nearest Neighbors (k-NN) algorithm

mgr inż. Paweł DRZEWIECKI

Autor jest doktorantem Śląskiego Środowiskowego Studium Doktoranckiego z dyscypliny Inżynieria Środowiska.



W KILKU SŁOWACH

W obecnych czasach ludzki intelekt zaczyna zastępować sztuczna inteligencja i związane z nią metody predykcji, czyli przewidywania, na podstawie zgromadzonych wcześniej danych. Nie ma firmy, która by nie starała się przewidzieć swoich potencjalnych zysków np na podstawie różnych danych statystycznych ustalając jednocześnie, który czynnik ma największy wpływ na rozwój firmy. Z tych przyczyn również sztuczna inteligencja znalazła zastosowanie w różnych dziedzinach przemysłu, gdzie nawet najdrobniejsza poprawa parametrów procesu może powodować znaczne obniżenie kosztów produkcji i wpłynąć bezpośrednio na wzrost konkurencyjności firmy na rynku.

Niniejsze opracowanie ma na celu przedstawienie metod sztucznej inteligencji pozwalające na analizę wpływu parametrów procesu stalowniczego na stopień aktywności tlenu, przy użyciu regresji liniowej, nieliniowej i nieparametrycznej a następnie na wyborze najlepszego modelu a następnie porównanie ich.



ABSTRACT

More and more often nowadays, for prediction we use artificial intelligence (AI) and its methods of forecasting based on existing data rather than our intellect. There is no company that has not attempted to predict their prospective profits based on information, such as statistical data, and determine the one factor that is crucial to the future of the company. This is why AI has found its use in various branches of the industry, where even the slightest improvement in process parameters can significantly reduce the costs of production and directly influence the market power of the company. The following study presents different AI strategies—ones that employ linear, non-linear and nonparametric regression—with which one can analyze the influence of steelmaking process parameters on the oxygen activity. By comparing these strategies it is possible to determine which one is the best.

Wstęp

Możliwość przewidywania zjawisk od wieków była przedmiotem zainteresowania człowieka. W tym celu, na przykład gromadzono latami

wiedze kilku pokoleń, aby wykorzystując ludzką inteligencję przewidzieć na przykład pogodę albo proste procesy technologiczne np. obróbkę cieplną mieczy dalmaceńskich.

Celem niniejszego artykułu jest przedstawienie algorytmu pozwalającego na zbadanie wpływu wartości parametrów mierzonych w piecu łukowym (EAF) na stopień aktywności tlenu w kąpeli stalowniczej za pomocą modelu inteligencji obliczeniowej w zagadnieniach predykcji wybranych parametrów procesu stalowniczego.

Proces optymalizacji sprowadza się zwykle do wykorzystania algorytmów gradientowych, które dążą do minimalizacji funkcji celu, zdefiniowanej w postaci błędu średniokwadratowego.

Zakres prowadzonych badań obejmował analizę wpływu parametrów procesu stalowniczego na stopień aktywności tlenu, przy użyciu regresji liniowej, nieliniowej i nieparametrycznej a następnie na wyborze najlepszego modelu a następnie porównanie ich.

Proces elektrostalowniczy w piecu elektryczno-łukowym

W niniejszym opracowaniu zostanie omówiony dokładnie drugi sposób otrzymywania stali (proces elektrostalowniczy) gdyż na jego podstawie zostały zebrane ogromne dane, które będą analizowane w dalszej części pracy. Ten rodzaj procesu wytapiania stali oparty jest na złomie stalowym, najczęściej zbiórkowym lub poamortyzacyjnym. Odgrywa to bardzo ważną

rolę w recyklingu stali. Procesy elektryczne służą do otrzymywania stali wysokiej, jakości. W procesach tych ciepło uzyskuje się z przemiany energii elektrycznej w ciepłą. Początkowo proces elektrostalowniczy w piecu elektryczno-łukowym (EAF) wykorzystywany był głównie do produkcji stali stopowych oraz jakościowych. Stosowane były wtedy głównie piece o małej mocy i pojemności. Z biegiem czasu rosły jednak rozmiary i moce pieców, co miało bezpośredni przebieg na masową produkcję stali tym sposobem[7].

Wyróżniamy następujące rodzaje pieców elektrycznych:

Oporowe – przetworzenie energii w czasie przepływu prądu przez wsad metaliczny albo w specjalnych elementach grzewczych, przekazyjących ciepło do nagrzewania metalu drogą promieniowania lub konwekcji,

Łukowe- przetworzenie energii w łuku elektrycznym :

- łuk pośredni (nad wsadem między poziomymi elektrodami),
- łuk bezpośredni (między poziomymi elektrodami),
- łuk zakryty (między zanurzonymi we wsadzie końcami elektrod),

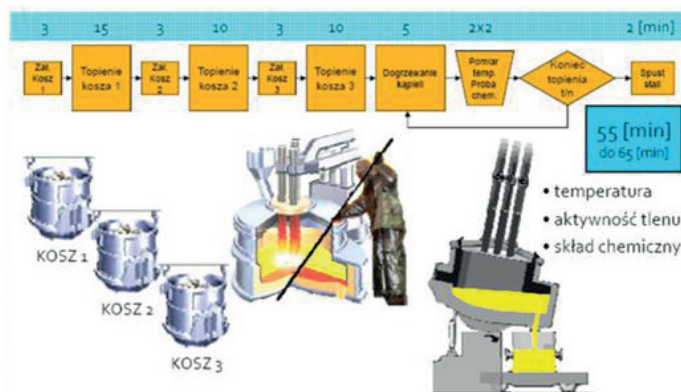
Indukcyjne – gdzie metal nagrzewa się prądem powstającym wskutek indukcji elektromagnetycznej :

- bezrdzeniowe o dużej częstotliwości (cewkę wtórną stanowi metal znajdujący się w tyglu,
- rdzeniowe o małej częstotliwości (cewkę wtórną stanowi pierścieniowa rynna pieca nad kąpielą metalową).

cza rosnąć będzie w podobnym tempie jak dotychczas.[7]

Ze względu na rodzaj pieca, w którym stal jest wytapiana, i wynikające stąd cechy charakterystyczne, procesy stalownicze można podzielić na [1]:

- martenowski (historyczny),
- proces konwertorowy, który można podzielić na poszczególne rodzaje:
 - z górnym dmuchem,
 - z dolnym dmuchem,
 - bessemerowski,
 - tomasowski,
 - z bocznym dmuchem,
 - kombinowany,
- proces elektrometalurgiczny
 - piece oporowe,
 - piece łukowe, które dzielą się na:
 - z łukiem pośrednim,
 - z łukiem bezpośrednim,
 - z zakrytym łukiem,
 - Piece indukcyjne:
 - bezrdzeniowe,
 - rdzeniowe.



Rys.1 Procesy w piecu EAF.[9]

Proces technologiczny wytwarzania stali

Wydawało się, iż stal zostanie wyparta przez nowoczesne materiały, do których głównie zaliczamy tworzywa sztuczne oraz wszelkiego rodzaju materiały kompozytowe, również o osnowie metalowej. Jednak okazało się, iż jako materiał konstrukcyjny stal pozostaje jednak niezastąpiona i ciągle kreowane są nowe obszary zastosowań wyrobów stalowych. Także ciągle doskonalony jest proces technologiczny produkcji stali. W ostatnich latach produkcja stali na świecie dynamicznie rośnie, a prognozy mówią, że również w latach następnych produkcja hutnicza

Zastosowanie tlenu w procesie elektropiecowniczym.

Na skalę światową, elektryczne piece łukowe do produkcji stali zaczęły być szerzej stosowane w ostatniej dekadzie. Wraz ze zwiększonym zastosowaniem paliw kopalnych oraz zanieczyszczonego złomu metali, wykorzystanie energii cieplnej i chemicznej tlenku węgla zawartego w spalinach zyskało rosnące znaczenie. Zależnie od warunków pracy pieca, do 30% energii doprowadzanej do pieca łukowego może zostać utracone w spalinach. Coraz częściej firmy proponują opatentowany system lancy tlenowej do



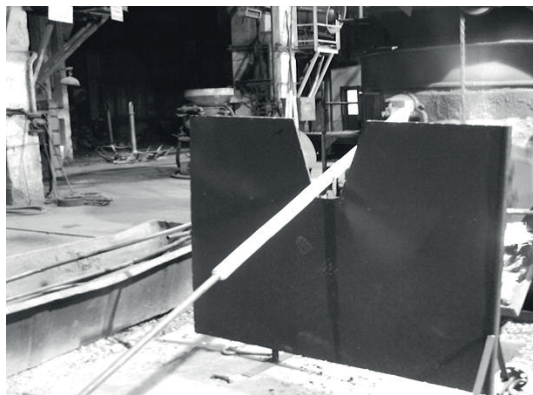


dopalania tlenku węgla w elektrycznych piecach łukowych (EAF). Poprzez selektywne zastosowanie tlenu wydajność instalacji produkcyjnych jest podwyższana, podczas gdy jednocześnie zużycie energii elektrycznej zostaje obniżone. Dla określenia docelowego zużycia tlenu konieczna jest znajomość dokładnego składu chemicznego spalin na wyjściu z pieca.

Wykorzystanie tlenu w hutnictwie dotyczy przede wszystkim [7]:

- Zastąpienia wdmuchanego powietrza dmuchem czystym tlenem lub powietrzem wzbogaconym w tlen (spalanie w tlenie).
- Podawania tlenu do lub nad płynną surówkę w celu usunięcia węgla oraz dopalenia CO.
- Przedmuchiwanie ciekłych metali tlenem w celu ich oczyszczenia oraz wymieszania prowadzącego do homogenizacji składu i temperatury.

Stosowanie tlenu zamiast powietrza w procesach spalania pozwala na zwiększenie wydajności paleniska, zmniejszenie straty kominowej poprzez ograniczenie ilości spalin o około 80 %, poprawę warunków wymiany ciepła, ograniczenie emisji tlenków azotu. Wszystkie wymienione korzyści wynikają przede wszystkim z faktu, że przez palenisko nie przepływa strumień azotu, który przy zasilaniu palników powietrzem przekracza prawie czterokrotnie strumień tlenu.



Rys. 2 Lanca tlenowa zabezpieczona rurkami szamotowymi, spoczywająca na osłonie z blachy stalowej.

Obszar stosowania inteligencji obliczeniowej

Obszarów stosowania eksploracji danych i inteligencji obliczeniowej jest wiele, obejmują one te miejsca, w których stosuje się systemy informatyczne, między innymi w celu gromadzenia pozyskanych danych w postaci baz da-

nych. Jesteśmy świadkami prawdziwej eksplozji baz danych, mając na myśli ich liczbę i objętość. Ogromne zbiory danych gromadzone są w hurtowniach danych. Z powodu dużej prostoty konstruowania bazy danych oraz akceptowalnych cen, systemy gromadzące dane stosuje się prawie we wszystkich dziedzinach życia. Wszędzie tam natomiast, gdzie istnieje już baza danych, pojawia się potrzeba analizy tych danych w celu odkrycia nieznannej dotąd wiedzy.

Eksplozja danych znajduje zastosowanie w wielu dziedzinach nauki stając się istotnym czynnikiem wpływającym na ich rozwój. Do obszarów zastosowań eksploracji danych należą [3]:

Medycyna – Zagadnienie eksploracji danych w medycynie jest chętnie wykorzystywane, tym bardziej, że wpływa na zwiększenie skuteczności leczenia, czy zapobiegania rozwoju groźnych chorób.

Przykład: Zastosowanie niektórych metod pozwala na tworzenie systemów wspierających diagnostykę, medyczną (jednym z przykładów zastosowania może być diagnoza płuc w celu wykrycia zmian nowotworowych).

Biznes – jest to dziedzina, w której eksploracja danych znalazła wiele praktycznych zastosowań, a umiejętnie stosowana może doprowadzić do zwiększenia zysków oraz dynamiki rozwoju przedsiębiorstwa.

Przykład: Bardzo przydatną cechą nowoczesnych systemów wspierających zarządzanie relacji z klientami (Relationship Management, CRM) jest możliwość dokonania skupień grup klientów, czy dokonywania analizy koszykowej w celu usprawnienia działalności marketingowej przedsiębiorstwa.

Technika – podobnie jak w przypadku wyżej wymienionych obszarów zastosowań metody eksploracji danych znalazły swoje zastosowanie w wielu gałęziach techniki.

Przykład: Diagnoza stanu skomplikowanej infrastruktury technicznej (np. sieci Eksploracja baz instalacji energetycznych czy pojedynczych jej elementów).

Biotechnologia – ze względu na dużą ilość danych występujących w badaniach biotechnologicznych. Eksploracja danych jest bardzo skutecznym narzędziem wspierającym badania w zakresie tej dziedziny.

Przykład: eksploracja danych w biotechnologii jest pomocna przy analizie zależności między różnymi strukturami DNA a prawdopodobieństwem wystąpienia raka.[8]



rys. 3 metody regresji

Metody parametryczne regresji

W obecnych czasach żadna firma nie mogłaby się rozwijać bez prognozowania przyszłości. Dotyczy to zarówno sprzedaży jak i samych procesów np. produkcyjnych w firmie. W tym celu wykorzystuje się różne metody regresji które pozwalają na badanie związku pomiędzy wielkościami danych i prognozowanie na tej podstawie nieznanych wartości jednych wielkości na podstawie znanych wartości innych. Formalnie regresja to dowolna metoda statystyczna pozwalająca estymować warunków wartość oczekiwaną zmiennej losowej, zwanej, dla zadanych wartości innej zmiennej lub wektora zmiennych losowych.[4]

Modele liniowe: regresja, analiza kowariancji i analiza wariancji

Z modelem mamy doczynienia. Jeśli pomiędzy dwiema wielkościami fizycznymi występuje zależność liniowa (a niemal zawsze może tak być, wystarczy tylko odpowiednio dobrać osie wykresu). Regresja liniowa jest prostą, (choć pracochłonną) metodą wyznaczenia parametrów najlepiej dopasowanej prostej. Uzyskane parametry dopasowania mogą następnie służyć do wyznaczenia szukanej wielkości fizycznej. [2]

Analiza wariancji jest jedną z najbardziej popularnych i najczęściej stosowanych analiz statystycznych. Dokładniej analizą wariancji określa się grupę analiz, służących do badania wpływu czynników (zmiennych niezależnych) na zmienną zależną.

Możemy podzielić analizę wariancji na trzy grupy analiz:

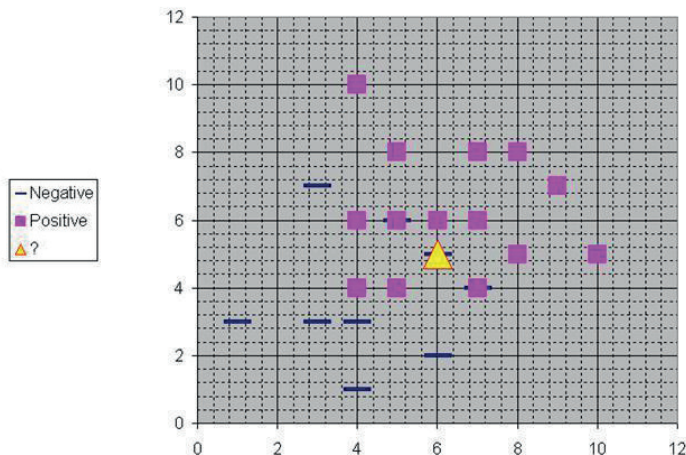
- Jednoczynnikowa analiza wariancji - wpływ jednego czynnika międzygrupowego na zmienną zależną
- Wieloczynnikowa analiza wariancji - wpływ kilku czynników międzygrupowych na zmienną zależną
- Analiza wariancji dla czynników wewnątrzgrupowych - wpływ czynnika wewnątrzgrupowego na zmienną zależną, tzw. "powtarzane pomiary".

Zdarza się również, że łączymy ze sobą różne rodzaje: międzygrupową (jedno bądź wieloczynnikową) z wewnątrzgrupową, co nazywa się mianem analizy wariancji w schemacie mieszanym. Ideą analizy wariancji jest sprawdzenie czy pewne zmienne niezależne (czynniki) mają wpływ na poziom zmiennej zależnej (zmiennej testowanej, mierzonej). W zależności od rodzaju czynników stosujemy różne rodzaje analizy wariancji. Analiza wariancji to stosunek wariancji, którą obliczyliśmy pomiędzy badanymi grupami a średnią wariancją, którą zaobserwowaliśmy wewnątrz grup (nie mylić z czynnikami między i wewnątrzgrupowymi)

Metoda k-NN

Jednym z algorytmów regresji nieparametrycznej używanych w statystyce do prognozowania wartości pewnej zmiennej losowej jest metoda k-NN, która może również być używana do klasyfikacji. Metoda k-najbliższych sąsiadów (k-NN) jest to metoda, w której zamiast dopasowywać model, wyszukujemy podobne obiekty. Metody takie nazywane są pamięciowymi (memory-based) lub analogowymi. Podstawą tej metody jest intuicyjne przeświadczenie, że podobne obiekty trafiają do tej samej klasy. Przewidywania metody k-najbliższych sąsiadów wyznaczanych są na podstawie k obiektów z próby uczącej, które są najbardziej podobne do obiektu, dla którego wyznaczamy wartość zmiennej zależnej. W przypadku zadań klasyfikacyjnych wykorzystywane jest głosowanie (voting), a dla problemów regresyjnych uśrednianie odpowiedzi dla k obiektów. [5]





Rys.4 Sposób klasyfikowania pojedynczych danych metodą k-NN.

Założenia:

☒ Dany jest zbiór uczący zawierający obserwacje, z których każda ma przypisany wektor zmiennych objaśniających oraz wartość d zmiennej objaśnianej Y.

☒ Dana jest obserwacja C z przypisanym wektorem zmiennych objaśniających dla której chcemy prognozować wartość zmiennej objaśnianej Y.

Przebieg algorytmu k-NN:

1. Porównanie wartości zmiennych objaśniających dla obserwacji C z wartościami tych zmiennych dla każdej obserwacji w zbiorze uczącym.
2. Wybór k (ustalona z góry liczba) najbliższych do C obserwacji ze zbioru uczącego.
3. Uśrednienie wartości zmiennej objaśnianej dla wybranych obserwacji, w wyniku czego uzyskujemy prognozę.

Sposób Doboru parametru k - liczby sąsiadów:

- Jeśli k jest małe, algorytm nie jest odporny na szumy – jakość klasyfikacji jest niska.
- Jeśli k jest duże, czas działania algorytmu rośnie - większa złożoność obliczeniowa.

Badania eksperymentalne

Bania polegały na przetestowaniu trzech baz danych (baza danych stworzona bezpośrednio z pomiarów i dwie po odpowiedniej obróbce wyjściowej) przy pomocy dwóch metod regresji: najbliższych sąsiadów (k-NN), oraz regresji liniowej.

7.1 Zbiór danych

Dane eksperymentalne pochodziły z bezpośrednich pomiarów dokonanych na jednym z piecu w polskiej hucie. Są to dane z kilku tysięcy wytopów, zarejestrowanych w przeciągu kilku lat. Proces ten składa się z dwóch etapów. Pierwsza część danych pochodzi z pierwszego etapu, w którym w elektrycznym piecu łukowym (EAF) topi się złom za pomocą łuku elektrycznego. W momencie, w którym temperatura oraz skład chemiczny stali są odpowiednie, stal jest zlewana do pieco-kadzi. Jednak przed rozpoczęciem drugiego etapu potrzebne jest odtlenienie stali. W związku z tym do stali wprowadzane są odtleniacze albo bezpośrednio przed zlanie stali odtleniacze są umieszczane w kadzi, drugim sposobem jest dodawanie odtleniaczy do strumienia stali zlewanej do pieco-kadzi.

Tak otrzymany zbiór danych został odpowiednio przekształcony przed przystąpieniem do eksperymentu. W sumie do badań użyto trzy bazy danych odpowiednio przygotowane, aby uzyskać jak najlepsze wyniki eksperymentów:

- 1) Zbiór Normal - to parametry bezpośrednio logowane z czujników.
- 2) Zbiór Sub - powstał w ten sposób że dla wybranych atrybutów które zmieniają się w trakcie procesu, jak np. "Temperatura Czujników TbX" wyznaczono zmianę temperatury pomiędzy kolejnymi pomiarami, czyli obliczono $TempTb1 - TempTb1Old$ gdzie $TempTb1$ - to aktualny pomiar temperatury, a $TempTb1Old$ to wartość temperatury podczas dokonania poprzedniego pomiaru. Tak zrobiono dla wszystkich zmieniających się w czasie pojedynczego procesu parametrów. Uzyskano w ten sposób różnice między kolejnymi pomiarami.
- 3) Zbiór Div - to zbiór w którym dodatkowo oprócz różnicy wartości ze zbioru Sub, uzyskany wynik różnicy podzielono przez czas w którym tą różnicę zaobserwowano, czyli np. $(TempTb1 - TempTb1Old)/(Czas - CzasOld)$, gdzie $Czas$ - to chwila czasowa w której dokonano ostatniego pomiaru, a $CzasOld$ to chwila czasowa w której dokonano poprzedniego pomiaru. Uzyskano w ten sposób różni-



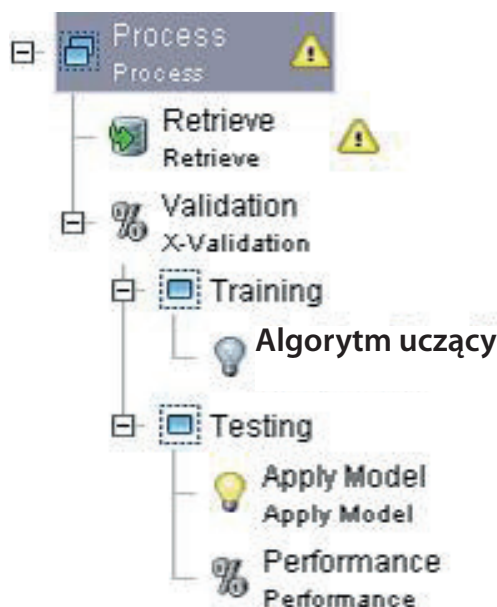
ce pomiarów ale w zależności od czasu między kolejnymi pomiarami.

Tak przygotowane dane miały na celu sprawdzić czy odstęp czasu między kolejnymi pomiarami ma wpływ na wyniki eksperymentów, co zostało przebadane w części doświadczalnej z wykorzystaniem narzędzia Rapid Miner.

Rapid Miner

Rapid Miner [4] jest złożonym środowiskiem służącym do prowadzenia doświadczeń i analiz z dziedziny uczenia maszynowego (ang. Machine learning) oraz eksploracji danych. Aplikacja ta udostępnia ponad 500 operatorów (z czego 100 z aplikacji Weka) dla wszystkich głównych procedur maszynowego uczenia, włączając w to obsługę wejścia, wyjścia, przetwarzania danych oraz wizualizację. Dzięki zastosowaniu XML-a i wygodnego edytora bez problemu można tworzyć łańcuchy i drzewa operatorów, źródeł danych i innych "bloków" składających się na program wykonywany przez Rapid Minera. Program ten wyróżnia, od innych narzędzi tego typu, czytelny interfejs użytkownika. Przyjazne ikony, dużo opisów, wszechobecne kreatory nie pozwalają użytkownikowi zagubić się w bogactwie opcji.

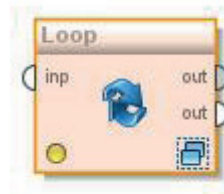
Schemat procesu



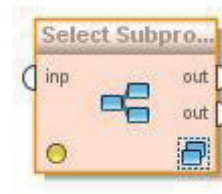
Rys. 5 Schemat blokowy działania algorytmu metody obliczeniowej.

W skład głównego procesu obliczeniowego dla danej bazy danych (operatora Loop parameters), wchodzi natępujące operatory :

- **Loop**
- **Select subprocess** – operator odpowiedzialny za wybór odpowiedniego procesu.



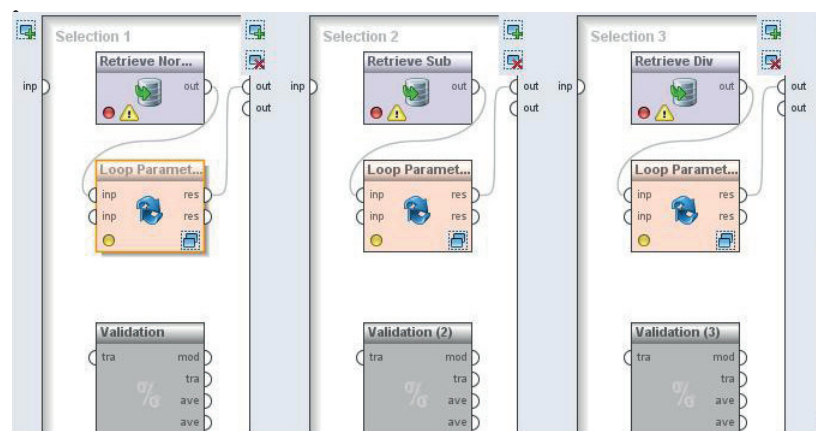
Rys.6 Główne okno procesu



Rys.7 Operator odpowiedzialny za wybór odpowiedniego zbioru danych.

W skład operatora Select Subprocess wchodzi następujące operatory (rysunek 8):

- **RTierve** – jest to operator umożliwiający odczytanie plików (Attribute-Relation File Format). Jest to plik tekstowy definiujący zbiory atrybutów, oraz konkretne instancje opisane tymi atrybutami. Został on opracowany na Wydziale Informatyki Uniwersytetu Waikato, do przeprowadzania uczenia maszynowego w programie Weka. Nagłówek pliku w formacie ARFF zawiera nazwę relacji (w przypadku programu RapidMiner – informacje te są ignorowane), oraz listę atrybutów. W tym przypadku jest on odpowiedzialny za wczytanie trzech badanych baz danych.
- **Loop Parameters** – operator odpowiedzialny za przejście procesu przez wszystkie zdefiniowane dla niego kombinacje parametrów.

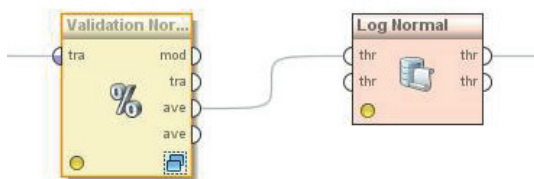


Rys.8 Procesy w Select Subprocess



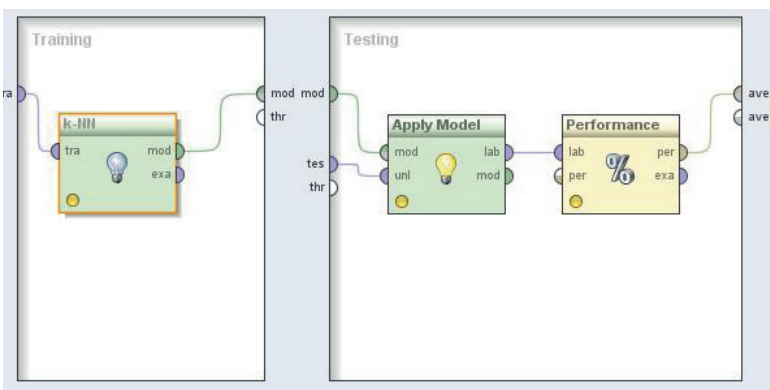
W skład operatora Loop Parameters wchodzi następujące operatory:

- Validation – Jest to operator testu krzyżowego. W przypadku badań wykorzystano 10-krotną walidację
- Log – operator odpowiedzialny za logowanie (zapisywanie) wyników z poszczególnych operatorów.



Rys. 9 Procesy w Loop Parameter

Rysunek Testowania klasyfikatora metodą cross-validation z wykorzystaniem k-NN i svm

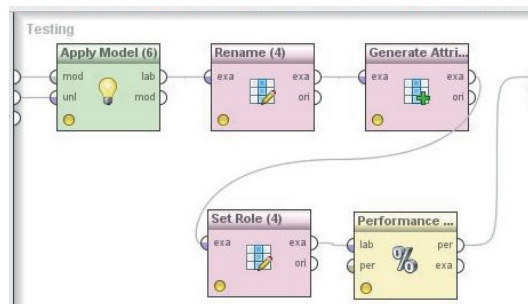


Rys. 10 Schemat doboru klasyfikatora

W skład operatora krosvalidacji wchodzi następujące operatory (dla bazy danych Nominal):

- Group Models – operator odpowiedzialny za pogrupowanie podłączonych modeli. Podczas prowadzonych badań były to dane pochodzące z dyskretyzacji oraz z procesu klasyfikacji (po dyskretyzacji).
- Apply model – operator ten umieszcza otrzymany model danych w ExampleSet. Posiada on również informację pochodzącą z danych treningowych. Informacje te są wykorzystane do przewidywania nieznanych wartości etykiet, powielania transformacji powstałych podczas procesu uczenia lub przeprowadzania innych zmian.
- Performance – operator odpowiedzialny za ocenę wyników.
- Operatory odpowiedzialne za sposób regresji. Odpowiednio:

- k-NN- operator
- Linear regression
- Set Role – operator ten umożliwia zmianę roli danego atrybutu (zwykle, specjalne,etykieta itp.) W tym przypadku przypisywano poszczególnym atrybutom rolę etykiety.



Rys. 11 Część testowa operatora krosvalidacji dla baz danych: sud i div

Dodatkowo dla baz danych sud i div:

- Rename – operator który może być użyty do zmiany nazwy atrybutu
- Generate Attribute - operator odpowiedzialny za This operator constructs new user defined attributes from mathematical expressions

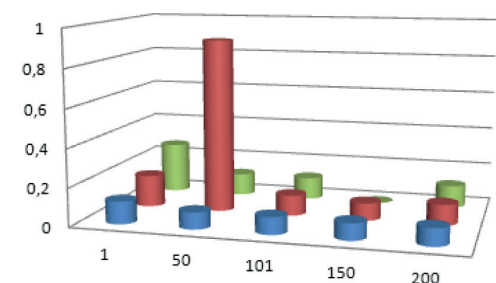
Wyniki dla metody k-NN

- Dla metody k-NN przeprowadzone zostały badania również dla trzech baz danych dla różnych parametrów k. Wyniki badań przedstawia poniższa tabela.

Lp.	Parametr k	Zbiór nominal		Zbiór Sud		Zbiór Div	
		(RMSE)	(RMSE) Std(+/-)	(RMSE)	(RMSE) Std(+/-)	(RMSE)	(RMSE) Std(+/-)
1	1	0,118	0,004	0,167	0,005	0,266	0,056
2	50	0,087	0,001	0,91	0,005	0,116	0,006
3	101	0,090	0,001	0,106	0,005	0,116	0,006
4	150	0,087	0,002	0,090	0,005	0,001	0,000
5	200	0,091	0,003	0,108	0,003	0,116	0,006

Tabela 1 Wynik testowania klasyfikatora metodą cross-validation z wykorzystaniem k-NN.

k-NN



■ Zbiór nominal
■ Zbiór Sud
■ Zbiór Div

Wykres 1 Rozkład wyników dla metody k-NN dla różnych wartości parametru k

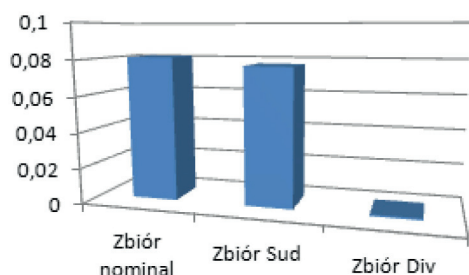
Wyniki Linear Regression

- Poniższa tabela przedstawia wyniki obliczeń przeprowadzonych metoda regresji liniowej dla 3 pomiarów. Z których następnie obliczono wartość średnią.

Lp.	Zbiór nominal		Zbiór Sud		Zbiór Div	
	(RMSE)	(RMSE) Std(+/-)	(RMSE)	(RMSE) Std(+/-)	(RMSE)	(RMSE) Std(+/-)
1	0,070	0,003	0,077	0,003	0,001	0,000
2	0,077	0,003	0,078	0,003	0,001	0,000
3	0,097	0,01	0,076	0,003	0,001	0,000
Średnia pomiarów	0,081		0,077		0,001	

Tabela 2 Wyniki pomiarów dla metody Linear Regression .

Regresja liniowa



■ Średnia pomiarów

Wykres 2 Wyniki pomiarów dla metody Linear Regression .

Podsumowanie

Liczba zgromadzonych informacji podwaja się, co około dwadzieścia miesięcy. Jednak rosnąca liczba informacji maszynowo odczytywalnych nie szła w parze ze zdolnością do ich zrozumienia i uczynie-

nia z nich użytku. Maszynowe uczenie dostarcza narzędzi, dzięki którym duże ilości danych mogą być automatycznie analizowane.

Zastosowanie w niniejszym artykule różnych metod inteligencji obliczeniowej miało na celu zweryfikowanie możliwości przewidywania sterowania podawaniem aktywnego tlenu. Przeprowadzone zostało to we współpracy z różnymi typami algorytmów regresyjnych, które charakteryzują się odmiennym sposobem podejmowania decyzji.

W pracy uwzględniono: metody predykcji takie jak: k-NN, regresję liniową, jako przykład prostego algorytmu regresji,

Metody te zostały przetestowane dla trzech odmiennych baz danych odpowiednio przygotowanych. Miało to na celu zbadanie wpływu obróbki zadanej bazy danych na jakość procesu.

Każda z tych metod realizuje odmienne sposoby zarówno uczenia się, jak i działania. Główny celem opracowania było przedstawienie sposobu postępowania przy przewidywaniu procesów przemysłowych oraz przybliżenie algorytmu postępowania .

LITERATURA

1. Adam Tabor;Metalurgia, Wydawnictwo Politechniki Krakowskiej, 1999
2. Siegmunt Brandt: Analiza Danych. PWN, 1999
3. David Hand: Eksploracja Danych: Wydawnictwa Naukowo Techniczne , 2005.
4. http://sztuczna-inteligencja.eprace.edu.pl/998,Narzedzia_Data_Mining.html,
5. Daniel T. Larose: Odkrywanie wiedzy z danych. .PWN, 2006
6. Daniel T. Larose: Metody i modele eksploracji danych. Wydawnictwo Naukowe PWN, 2008.
7. A. Chojecki: Procesy wytapiania stali na odlewy. Wyd. AGH, Kraków 1991.
8. Zdzisław Pawlak: Rough Sets. International Journal of Computer and Information Sciences, 11: 341-356, 1982
9. Błachnik M., Kordos M., Wieczorek T.: Temperature Prediction in Electric Arc Furnace with Neural Network Tree,

XV MIĘDZYNARODOWA KONFERENCJA NAUKOWA

Materiały ogniotrwałe:
wytwarzanie, metody badań, stosowanie

22–24.05.2013
WISŁA-JAWORNIK

Celem konferencji jest zaprezentowanie najnowszych osiągnięć inżynierii materiałowej w zakresie: technologii wytwarzania, stosowania i metodyki badań materiałów ogniotrwałych wykorzystywanych w procesach przemysłowych.

organizatorzy:



Instytut Ceramiki i Materiałów Budowlanych



IPMO



SITP

icimb.pl/gliwice/konferencja

32 270 19 22
32 270 18 56

mkh@icimb.pl

Konferencja jest dofinansowana ze środków Ministerstwa Nauki i Szkolnictwa Wyższego na działalność upowszechniającą naukę.