



METRO

MEtalurgiczny TRening *On-line*



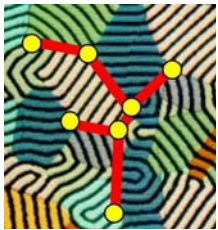
Komputerowe metody analizy i sterowania procesami produkcyjnymi

Marcin Perzyk

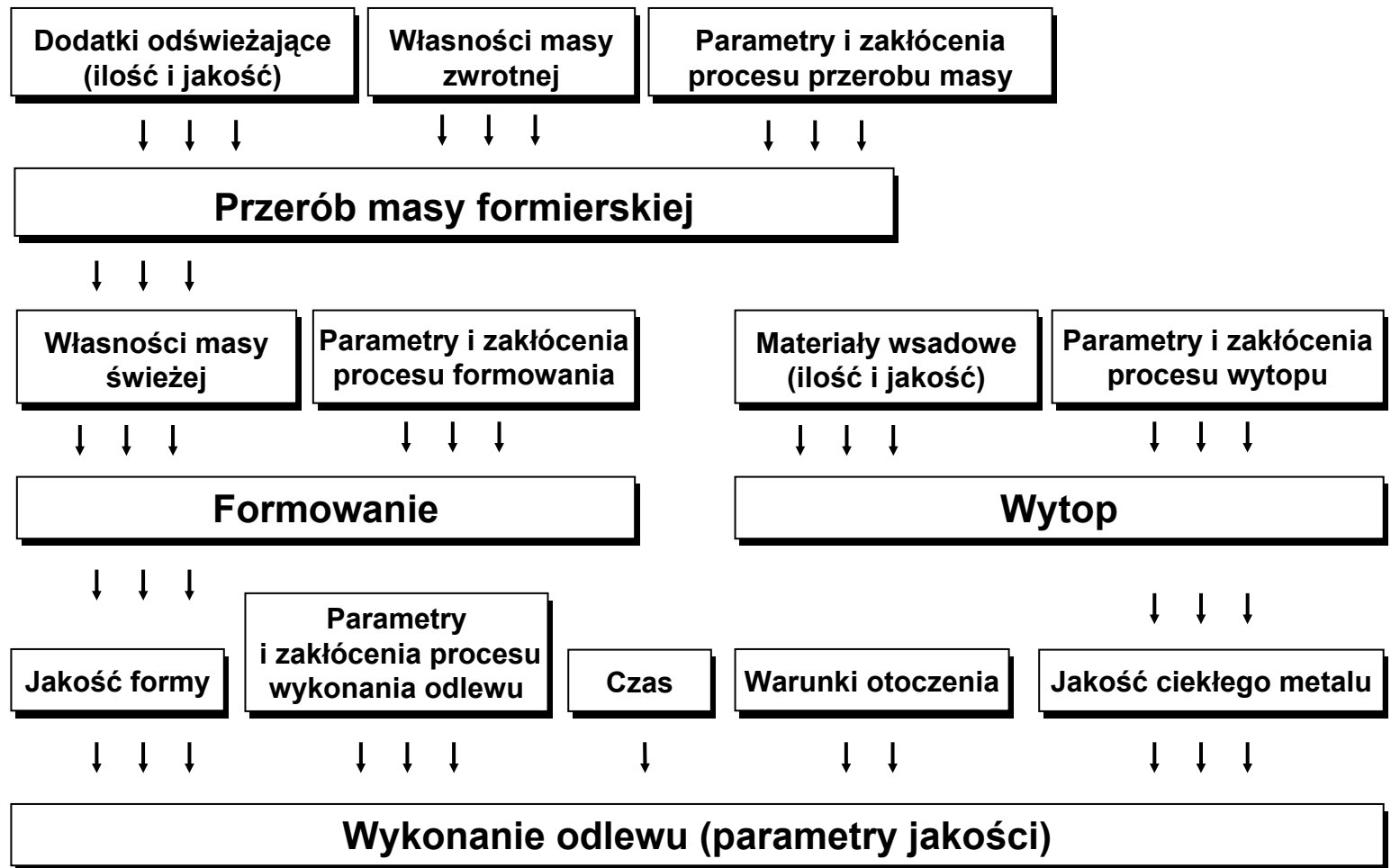
PW



Edukacja i Kultura



Główne procesy produkcyjne w typowej odlewni

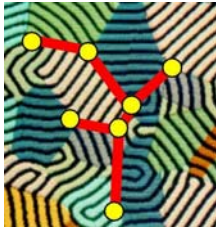




Typy zadań związanych procesami produkcyjnymi wspomaganych komputerowo



- Przewidywanie (symulacja) przebiegu procesu
- Bieżąca kontrola procesów i sterowanie nimi
- Eksploracja danych w celu wykrywania prawidłowości występujących w procesach



Przewidywanie (symulacja) przebiegu procesu



- Główne zastosowanie na etapie projektowania procesów (np. symulacja numeryczna wypełniania formy przy projektowaniu układu wlewowego)
- Umożliwia także przewidywanie skutków wprowadzenia zmian w technologii lub organizacji produkcji

Symulacja wymaga zastosowania modelu procesu



Bieżąca kontrola procesów produkcyjnych i sterowanie nimi



- Obejmuje *metody* statystycznego sterowania procesem (SSP) oraz *strategie* czy *metodologie*, np. typu SixSigma
- SSP pozwala na wykrywanie zakłóceń procesu oraz ocenę jakości procesu
- *Metodologie* wskazują na sposoby i organizację postępowania dla uzyskania i utrzymania właściwego przebiegu procesu
- Metody SSP *jedynie ułatwiają diagnozę* przyczyn rozregulowań procesu, którą dokonuje człowiek

Tematyka ta jest przedmiotem odrębnego wykładu

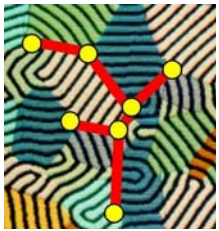


Wykrywanie prawidłowości występujących w procesach



- Wykrywanie przyczyn zakłóceń procesów (np. przyczyn powstawania braków wyrobów finalnych lub pośrednich)
- Umożliwia wskazanie optymalnych lub krytycznych parametrów procesu (np. kombinacji czasu i temperatury obróbki cieplnej)
- Stosuje nowoczesne metody typu eksploracja danych (data mining)

Wymaga zastosowania modelu procesu



Modelowanie procesów

Pojęcia podstawowe



- *Modelem* nazywamy uproszczony obiekt, który zachowuje się pod względem badanych zjawisk tak, jak obiekt rzeczywisty
- Rozróżnia się dwa *typy modeli*: fizyczne i matematyczne
- *Model fizyczny* wykorzystuje podobieństwo fizyczne między obiektem rzeczywistym, a modelem, stanowiącym pewne stanowisko doświadczalne. Zasady budowy takich modeli doświadczalnych formułuje teoria podobieństwa
- *Model matematyczny* stanowi równanie matematyczne lub układ równań lub innego typu związki opisujące dane zjawisko czy proces

Rozwój technik obliczeniowych sprawił, że dominuje modelowanie matematyczne (najczęściej numeryczne)



Modelowanie procesów

Rodzaje modeli matematycznych



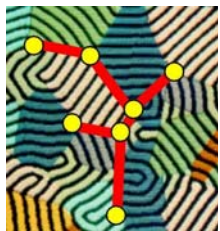
- Związki uwzględniające naturę danego zjawiska czy procesu (np. prawa fizyczne rządzące wymianą ciepła między odlewem a formą). Odnoszą się głównie do procesów powstawania odlewu w formie, tj. zalewania, krzepnięcia, powstawania naprężeń w odlewie.
Stanowią przedmiot innych wykładów kursu.
- Modele typu ‘czarna skrzynka’, nie uwzględniające natury zjawisk, stosowane do dowolnych procesów produkcyjnych.
Stanowią przedmiot niniejszego wykładu.



Modele nie uwzględniające natury zjawisk i procesów



- Stałe (parametry) modelu wyznaczone są na podstawie danych doświadczalnych, zebranych najczęściej w toku normalnego procesu produkcyjnego, rzadziej w specjalnie zaplanowanych eksperymentach.
- Stosuje się różnego rodzaju algorytmy i metody optymalizacyjne, zależnie od rodzaju modelu.
- W pełni określony model, tj. z wyznaczonymi wartościami parametrów może następnie być wykorzystany do przewidywania wartości wyjściowych (zmiennych zależnych) dla nowych wartości zmiennych wejściowych lub ich kombinacji, tj. takich, które nie wystąpiły w zbiorze użytym do wyznaczenia stałych modelu.



Główne rodzaje modeli nie uwzględniających natury zjawisk i procesów



- Typu statystycznego
- Wykorzystujące metody sztucznej inteligencji, w szczególności systemy uczące się (metody uczenia maszynowego)

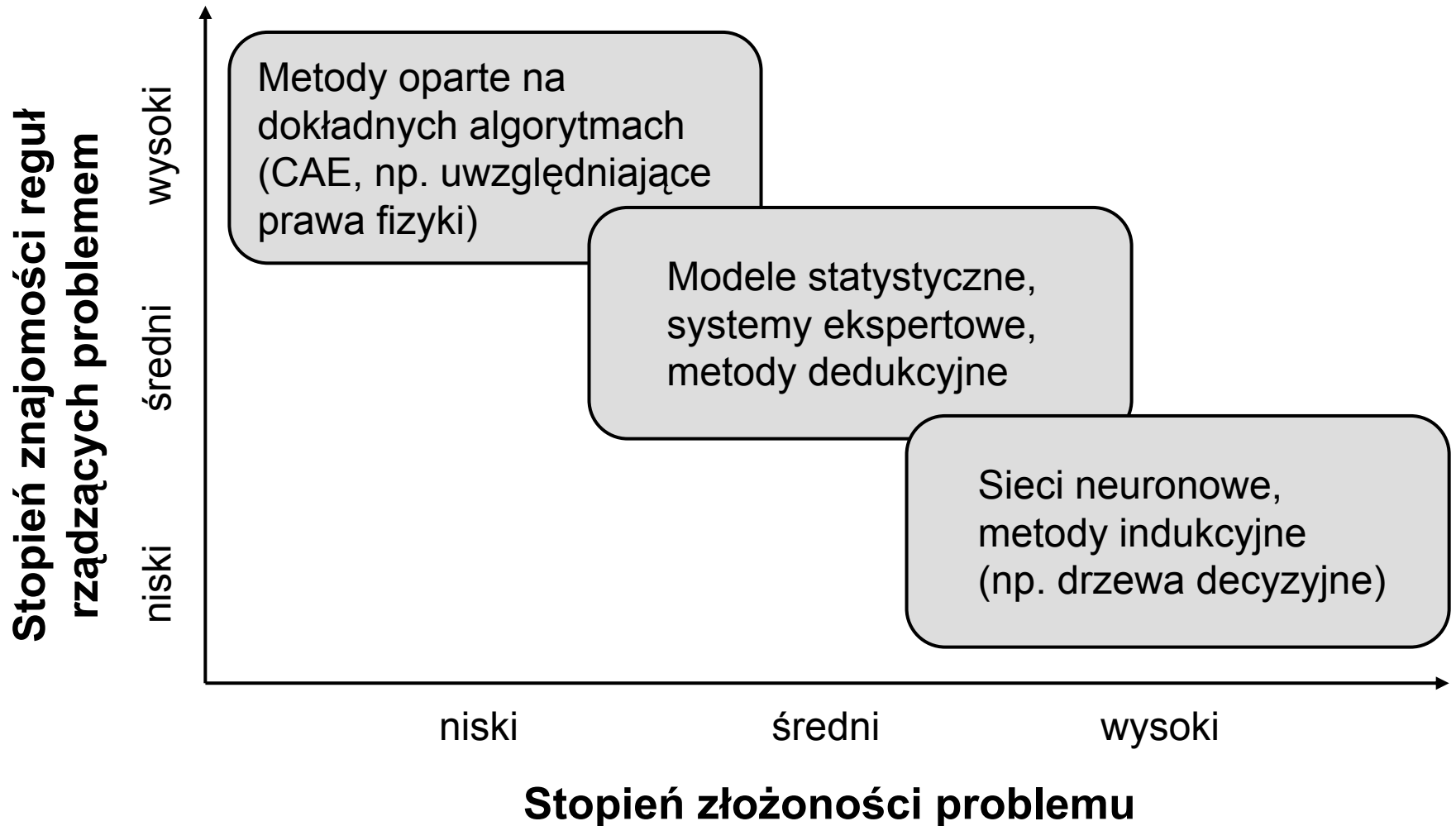
Definicja

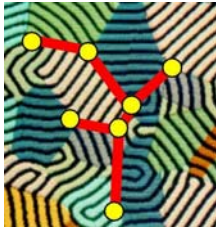
System uczący się to taki, w którym na podstawie doświadczeń (przykładów uczących) zachodzą zmiany prowadzące do poprawy jego działania



Zakres stosowania różnych typów modeli matematycznych

(wg R. Tadeusiewicza)





Typy danych występujących w modelowaniu procesów



Dane zawarte w przykładach uczących, zwane *atrybutami*, mogą być następujących typów:

- Nominalny, tj. o skończonym zbiorze nieuporządkowanych wartości dyskretnych, zwanych kategoriami.
Często wartości te określa się w sposób werbalny, np. piec topialny może być: *elektryczny oporowy*, *elektryczny indukcyjny*, *gazowy*; wyrób (odlew) może być *dobry* lub *wadliwy*.
- Porządkowy, tj., o przeliczalnym zbiorze uporządkowanych wartości dyskretnych.
Np. temperatura może być: *niska*, *średnia*, *wysoka*.
- Liczbowy ciągły, tj. o wartościach ze zbioru liczb rzeczywistych.



Typy danych uczących wykorzystywanych w modelowaniu procesów



Niektóre modele wymagają stosowania wielkości typu nominalnego lub porządkowego. Uwzględnianie wielkości typu ciągłego jest również możliwe przy zastosowaniu następującego postępowania:

- Zmienne ciągłe można zamieniać na wartości atrybutów wyrażone przez kategorie zaliczając daną zmienną do oznaczonego numerem odpowiedniego przedziału jej oryginalnej wartości (zamiana na dane typu porządkowego).
- W przypadku, gdy kategorie wyjścia zostały utworzone w powyższy sposób, możliwa jest zamiana nazwy obliczonej przez model kategorii (numeru) na wartość ciągłą, obliczoną np. jako średnia z takiego przedziału.

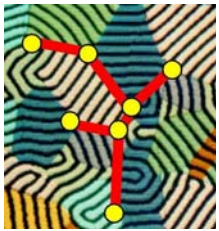


Typy danych uczących wykorzystywanych w modelowaniu procesów



Zamiana wielkości typu ciągłego na wartości typu nominalnego lub porządkowego powinna być dokonana z uwzględnieniem następujących uwarunkowań:

- Mała liczba kategorii (przedziałów) zmniejsza precyzję modelu, zarówno przy uczeniu się, jak odpytywaniu.
- Duża liczba kategorii stwarza z kolei niebezpieczeństwo, że niektóre z nich będą nie reprezentowane w zbiorze uczącym, a przynajmniej reprezentacje będą bardzo mało liczne.



Rodzaje zadań realizowanych najczęściej przez modele procesów produkcyjnych



- **Regresja (aproksymacja funkcji)**
Dopasowanie do zbioru punktów doświadczalnych zależności analitycznej (wyrażonej wzorem) między ciągłymi zmiennymi wejściowymi (niezależnymi), a ciągłą zmienną wyjściową (zależną). *Przykład:* wytrzymałość stopu w zależności od jego składu chemicznego.
- **Klasyfikacja**
Przypisywanie przypadków do jednej z określonej liczby klas (kategorii, czyli wartości) reprezentowanych w wyjściowej zmiennej nominalnej lub porządkowej. *Przykład:* klasyfikacja wyrobu jako dobry lub zły.
- **Wykrywanie regularności**
Wykrywanie istotnych cech w danych (obrazach) wejściowych, bez wiedzy na temat wzorców. *Przykład zastosowania:* grupowanie postaci konstrukcyjnych części maszyn.



Modele statystyczne

Informacje podstawowe



- Realizują zadania typu regresji (aproksymacji funkcji)
- Ogólna postać funkcji:

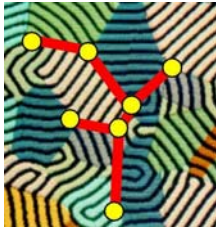
$$y = f(x_1, x_2, \dots, a_1, a_2, \dots)$$

x_i – zmienne niezależne,

y – zmienna zależna,

a_j – stałe (parametry funkcji), które należy wyznaczyć na podstawie danych doświadczalnych

- Konkretną postać funkcji należy założyć



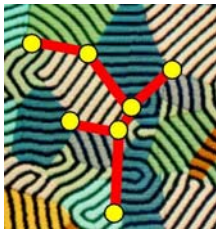
Modele statystyczne

Rodzaje stosowanych funkcji



- Jednej zmiennej lub wielu zmiennych (regresja jedno lub wielowymiarowa)
- Liniowe i nieliniowe (regresja liniowa i nieliniowa)
 - Funkcje nieliniowe:
 - wielomiany dowolnego stopnia
 - inne dowolne funkcje (np. potęgowe, wykładnicze itp.)

Dobór postaci funkcji przeprowadza się zwykle po naniesieniu danych doświadczalnych na wykres. Znaczne trudności występują w przypadku funkcji wielu zmiennych.



Modele statystyczne

Wyznaczanie parametrów modeli



- Kryterium *minimum błędu* ε modelu określanego jako suma kwadratów odchyleń od wartości obserwowanych:

$$\varepsilon = \sum_{k=1}^n (y_k - d_k)^2$$

gdzie n oznacza liczbę przykładów doświadczalnych, y_k – wartości przewidywane przez funkcję (zależne od parametrów a_i), zaś d_k – wartości wyjściowe w danych doświadczalnych (przykładach)

- Analityczne (jednoznaczne) sposoby wyznaczania parametrów równania są możliwe tylko dla postaci liniowej lub funkcji typu wielomianu dowolnego stopnia.
- Dla pozostałych przypadków stosuje się:
 - linearyzację funkcji (najczęściej), lub
 - zastosowanie innych metod optymalizacji parametrów (rzadko)



Modele statystyczne liniowe i w postaci wielomianów

Wyznaczanie parametrów

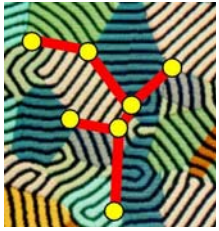


- Wykorzystuje się układ równań liniowych uzyskany po zróżniczkowaniu wartości błędu:

$$\left(f(x_1, x_2, \dots, a_1, a_2, \dots)_k - d_k \right)^2$$

względem, kolejno, poszczególnych parametrów a_i , przyrównaniu tak obliczonych pochodnych cząstkowych do zera i zsumowaniu tych równań, dla wszystkich k-tych przykładów danych.

- Otrzymany układ zawiera tyle równań, ile jest niewiadomych parametrów równania regresyjnego. Minimalna liczba przykładów, umożliwiająca wyznaczenie parametrów równa jest liczbie tych parametrów.



Modele statystyczne nieliniowe

Wyznaczanie parametrów



- Przykłady linearyzacji zależności przez wprowadzenie nowej zmiennej (X lub Y):

$$y = \frac{x}{a_1 \cdot x + a_2} \Rightarrow Y = a_1 \cdot x + a_2 \quad Y = \frac{x}{y}$$

$$y = a_1 \cdot x^{a_2} \Rightarrow Y = \log(a_1) + a_2 \cdot X \quad X = \log(x), \quad Y = \log(y)$$

- Inne metody optymalizacji (minimalizacji błędu modelu) omówione będą przy modelach typu sztucznych sieci neuronowych oraz algorytmach genetycznych



Najważniejsze modele procesów produkcyjnych wykorzystujące metody sztucznej inteligencji



- Sztuczne sieci neuronowe
Najczęściej stosowane i o największych możliwościach. Są przedmiotem odrębnego wykładu
- Drzewa decyzyjne
- Modele stosujące logikę lub rachunek liczb rozmytych
- Modele stosujące optymalizację genetyczną
- Klasyfikacja bayesowska
- MARSplines



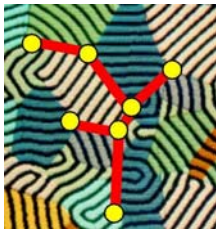
Drzewa decyzyjne

Informacje podstawowe



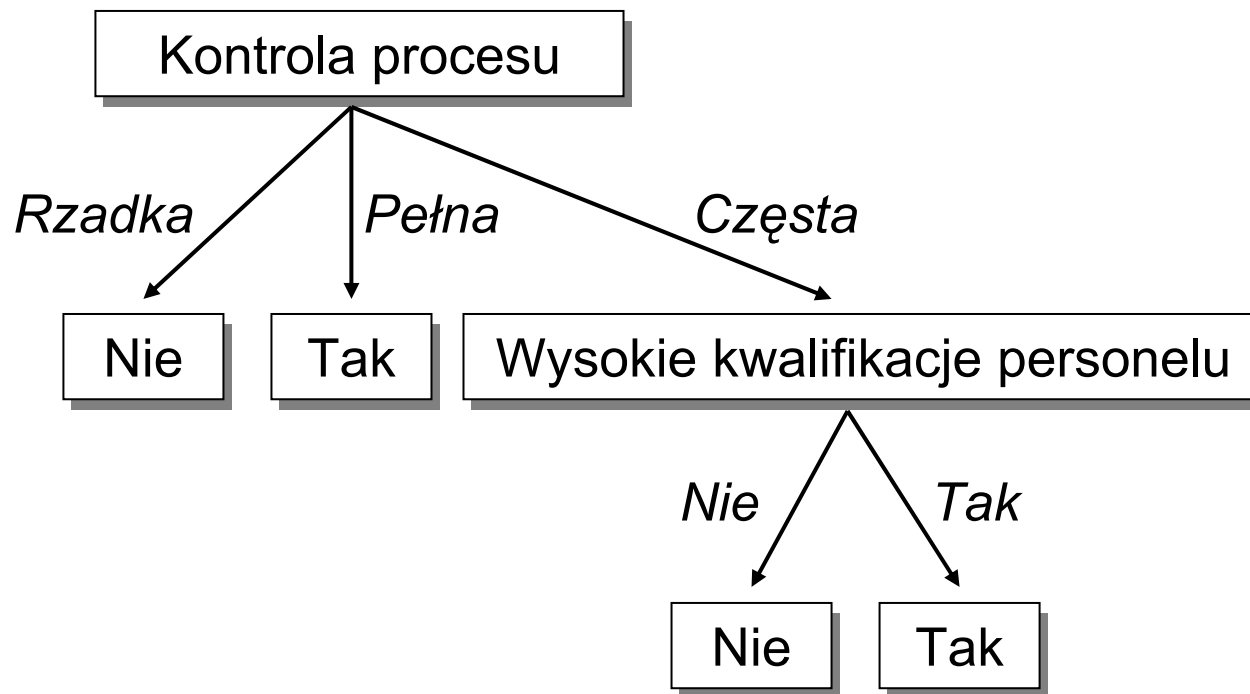
Drzewo decyzyjne jest to system operujący na danych typu nominalnego lub porządkowego. Jest strukturą logiczną (grafem) składającą się z elementów opisanych poniżej

- *Korzeń* (początek drzewa), z którego wychodzą co najmniej dwie gałęzie do *węzłów* leżących na niższym poziomie
- Z każdym węzłem związany jest *test* sprawdzający wartości atrybutów opisujących przykłady (uczące lub zadane, dla których chcemy znaleźć odpowiedź systemu).
- Dla każdego z możliwych wyników testu odpowiadająca mu gałąź prowadzi do węzła leżącego na niższym poziomie.
- Węzły, z których nie wychodzą żadne gałęzie są to *liście*, którym przypisane są klasy.

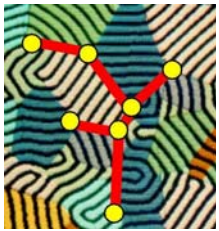


Drzewa decyzyjne

Ilustracja przykładowej struktury drzewa



Drzewo decyzyjne dla wielkości (pojęcia) „niski poziom braków”, uzależnionego od kilku zanotowanych parametrów produkcyjnych (wymienionych dalej w tabeli)



Drzewa decyzyjne

Dane uczące użyte do wygenerowania przykładowego drzewa



Nr przykładu	Kontrola procesu	Wysokie kwalifikacje personelu	Wielkość zakładu	Niski poziom braków
1	częsta	tak	duży	tak
2	częsta	nie	mały	nie
3	pełna	tak	mały	tak
4	rzadka	tak	duży	nie
5	rzadka	tak	mały	nie
6	częsta	tak	mały	tak
7	rzadka	nie	mały	nie
8	częsta	nie	duży	nie



Drzewa decyzyjne

Generowanie (indukcja)



- Do generowania drzew decyzyjnych na podstawie obserwacji stosuje się programy komputerowe stosujące różnego typu algorytmy
- Przykładowe drzewo wygenerowano z wykorzystaniem klasycznego algorytmu ID3
W wyniku działania programu okazało się m.in., że optymalnym atrybutem dla korzenia drzewa jest rodzaj kontroli procesu oraz że wielkość zakładu nie ma wpływu na poziom braków
- Inne, nowsze algorytmy generowania pozwalają m.in. na uwzględnienie danych niepełnych i niejednoznacznych, a także na upraszczanie zbyt złożonych drzew



Drzewa decyzyjne

Podstawowe rodzaje zadań (zastosowań)



- Klasycznym zastosowaniem drzew decyzyjnych jest klasyfikacja (drzewa działają na danych typu nominalnego lub porządkowego)
- Możliwa jest również realizacja zadań typu regresji (aproksymacji funkcji)

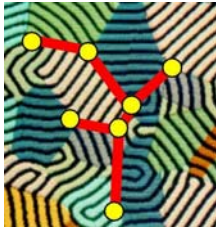
Uwzględnienie atrybutów o wartościach typu ciągłego wymaga przyjęcia granic przedziałów tych wartości dla testów w węzłach drzewa. Forma uzyskiwanych wyników (zazwyczaj w postaci nierówności) jest inna, niż w przypadku analizy regresji wykonanej innymi metodami, np. statystycznymi.



Drzewa decyzyjne ze wzmacnianiem (boosted trees)



- Dla trudnych zadań estymacji i predykcji, przewidywania generowane przez *sekwencje* prostych drzew są bliższe rzeczywistym wartościom, niż prognozy jednego, złożonego drzewa.
- Technikę polegającą na stosowaniu sekwencji prostych modeli, przy czym każdy kolejny model przykłada większą "wagę" do tych obserwacji, które zostały błędnie zaklasyfikowane przez poprzednie modele, nazywamy *wzmacnianiem* (ang. boosting).
- Drzewa decyzyjne ze wzmacnianiem lepiej modelują zależności złożone (w porównaniu do drzew prostych), ale są trudniejsze w interpretacji oraz wymagają większych nakładów obliczeniowych

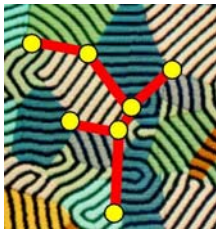


Modele stosujące logikę rozmytą i rachunek liczb rozmytych



Informacje podstawowe

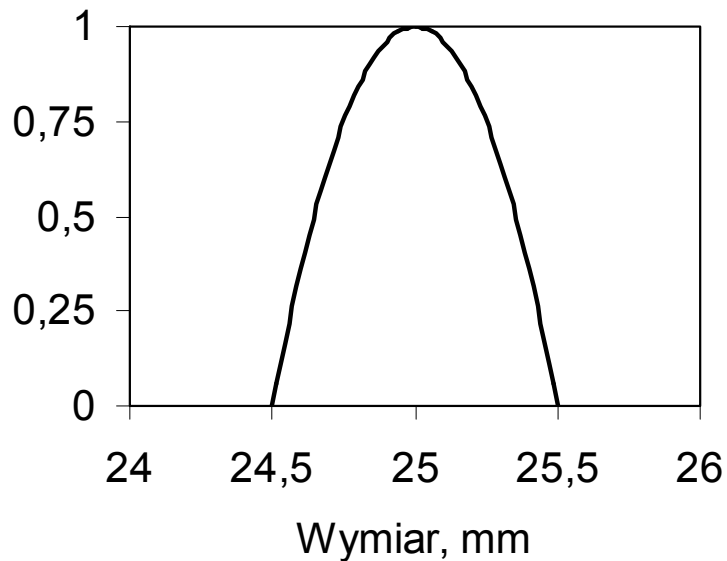
- W celu wyrażania wielkości określanych *nieprecyzyjnie*, w sposób przybliżony, wprowadzono zbiory i liczby rozmyte.
- Nieprecyzyjnie określona wielkość opisana jest nie tylko przez zakres możliwych wartości (np. wymiar „około 25 mm” może zawierać się w przedziale od 24,5 do 25,5 mm), ale także przez tzw. *funkcję przynależności* (do zbioru) lub inaczej *preferencji*, opisującą stopień pożądanego przyjęcia danej wartości z tego przedziału.
- Funkcja przynależności (preferencji) może przyjmować wartości z przedziału 0,1. Zbiór rozmyty zdefiniowany na osi liczb rzeczywistych o ciągłej i wypukłej funkcji przynależności nazywamy liczbą rozmytą.



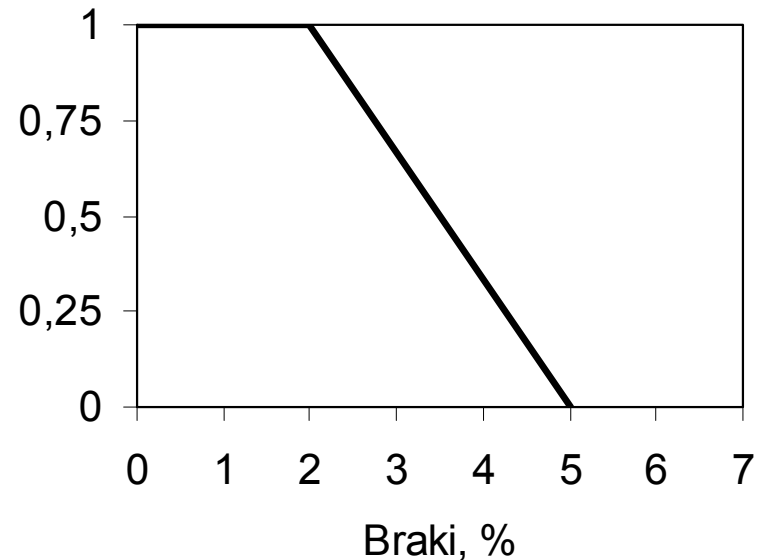
Modele stosujące logikę rozmytą i rachunek liczb rozmytych



Przykłady funkcji przynależności



Przykładowa funkcja przynależności (preferencji) dla wymiaru „około 25”



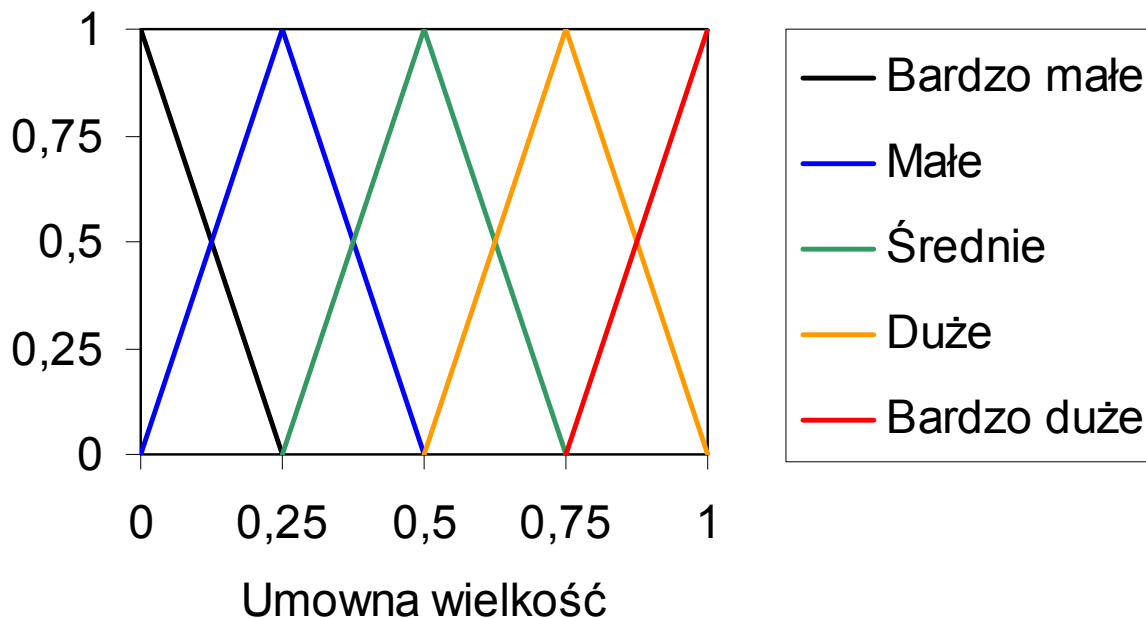
Przykładowa funkcja przynależności (preferencji) dla poziomu braków „niski”



Modele stosujące logikę rozmytą i rachunek liczb rozmytych



Przykłady funkcji przynależności



Często stosowane trójkątne funkcje przynależności (preferencji) dla nieprecyzyjnego określania wartości dowolnego parametru



Modele stosujące logikę rozmytą i rachunek liczb rozmytych



- Logika rozmyta jest nowoczesnym, znaczącym działem matematyki.
- Najczęstszym zastosowaniem logiki rozmytej (przybliżonego wnioskowania) są sterowniki rozmyte stosowane powszechnie w automatyce.
- Sterownik rozmyty wymaga stworzenia bazy reguł logiki rozmytej, np. jeżeli temperatura jest „wysoka” i wilgoć jest „średnia”, to nastawienie mocy klimatyzatora powinno być „wysokie”.
- Reguły takie można projektować lub tworzyć na podstawie liczbowych danych doświadczalnych.
- Reguły logiki rozmytej mogą być także tworzone z wykorzystaniem systemów uczących się, zwłaszcza sztucznych sieci neuronowych. Metoda ta wykorzystywana jest w budowie sterowników neuronowo – rozmytych, szeroko stosowanych we współczesnym przemyśle.



Modele stosujące logikę rozmytą i rachunek liczb rozmytych



- Liczby rozmyte mogą zastąpić wartości „zwykłe” w szeregu zależnościach analitycznych, tworząc modele bardziej realistyczne i dające nowe możliwości interpretacyjne.
- Wielkości określane nieprecyzyjnie występują technice bardzo często w różnych sytuacjach związanych np. z projektowaniem wyrobów i procesów.
- Pokazane zostanie zastosowanie liczb rozmytych do wyznaczania wskaźnika oceny wariantu projektowego.



Modele stosujące liczby rozmyte

Przykład zastosowania



- Klasyczny (ostry) wskaźnik oceny j -tej wersji projektowej procesu lub wyrobu definiuje się jako:

$$\gamma_j = \sum_{i=1}^n P_{ij} \cdot \alpha_i$$

gdzie P_{ij} – wartość i -tego parametru dla j -tej wersji projektowej, α_i – znaczenie i -tego parametru dla oceny

- Zamiast ostrych wartości do wzoru tego można podstawić liczby rozmyte, np. zdefiniowane werbalnie jako „bardzo małe”, „małe”, itd., którym odpowiadają funkcje przynależności np. jak na ostatnim wykresie.



Modele stosujące liczby rozmyte

Przykład zastosowania



Funkcja przynależności w postaci trójkąta zdefiniowana jest przez trójkę liczb określającą położenie granic przedziału l i p (wartości = 0) oraz maksimum m (wartość = 1). Dla takich funkcji działania na liczbach rozmytych typu iloczynu i sumy są proste:

- suma dwóch liczb rozmytych x i y jest liczbą rozmytą o kształcie trójkątnym, definiowaną trójką liczb wyznaczaną zgodnie z poniższym zapisem:

$$(l_x, m_x, p_x) \oplus (l_y, m_y, p_y) = (l_x + l_y, m_x + m_y, p_y + p_y)$$

- iloczyn takich dwóch liczb (z pewnym przybliżeniem):

$$(l_x, m_x, p_x) \otimes (l_y, m_y, p_y) \cong (l_x \cdot l_y, m_x \cdot m_y, p_y \cdot p_y)$$

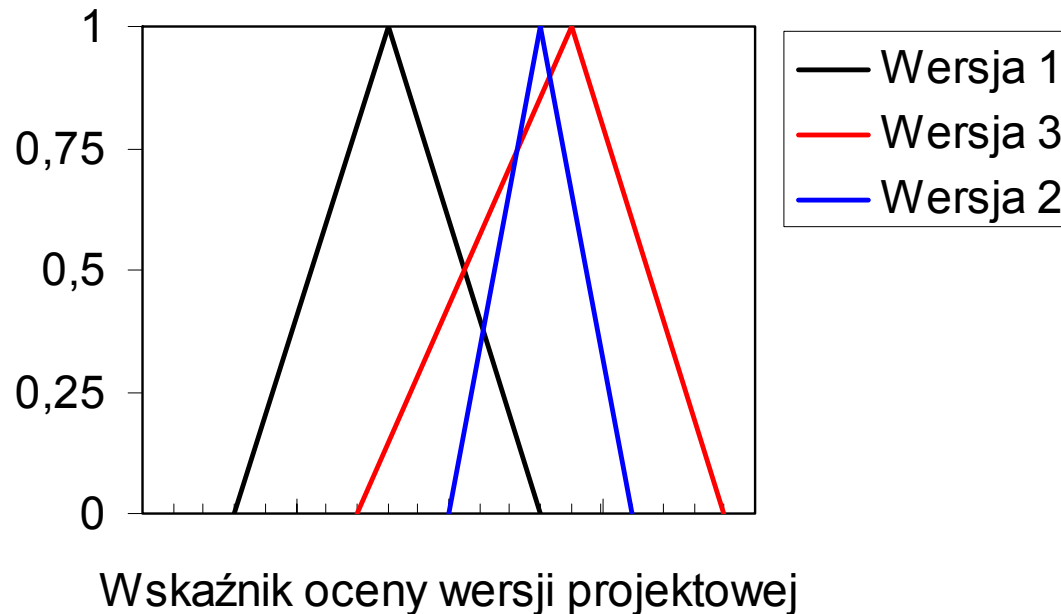


Modele stosujące liczby rozmyte

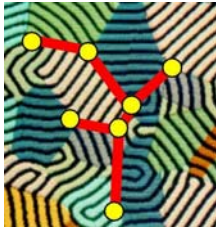
Przykład zastosowania



Przykładowe wyniki obliczeń rozmytych wskaźników oceny wersji projektowych.



Wersja 1 jest wyraźnie gorsza, natomiast decyzja, czy wybrać Wersję 2 (pewniejszą, ale z maksimum gorszym), czy Wersję 3 (z maksimum lepszym, ale mniej bardziej niepewną) nie jest jednoznaczna.



Modele stosujące optymalizację genetyczną



- Algorytmy genetyczne są nowoczesnym i efektywnym narzędziem matematycznym służącym do *optymalizacji dowolnej funkcji* jednej lub wielu zmiennych, wzorowanym na naturalnej ewolucji.
- Optymalizowana funkcja, nazywana *funkcją celu* lub *funkcją przystosowania*, stanowi model danego problemu. Przykładem może być minimalizowany czas przestojów linii produkcyjnej w funkcji parametrów charakteryzujących harmonogram wykonywania poszczególnych operacji.
- Funkcja celu może mieć dowolną postać, np. ciągłą typu wzoru analitycznego jak i dyskretną. Istotne jest tylko to, ażeby po podstawieniu wartości zmiennych niezależnych można było obliczyć jej wartość.



Modele stosujące optymalizację genetyczną

Cechy algorytmów genetycznych



- Algorytmy genetyczne nie przetwarzają bezpośrednio poszukiwanych parametrów modelu lecz ich zakodowaną postać, w postaci ciągów liczb binarnych (*genów*), zwanych *chromosomami*
- Przeszukiwanie wychodzi nie z jednego punktu ale z pewnej ich *populacji* (zbioru chromosomów), np.

chromosom A	11011001
chromosom B	10010010

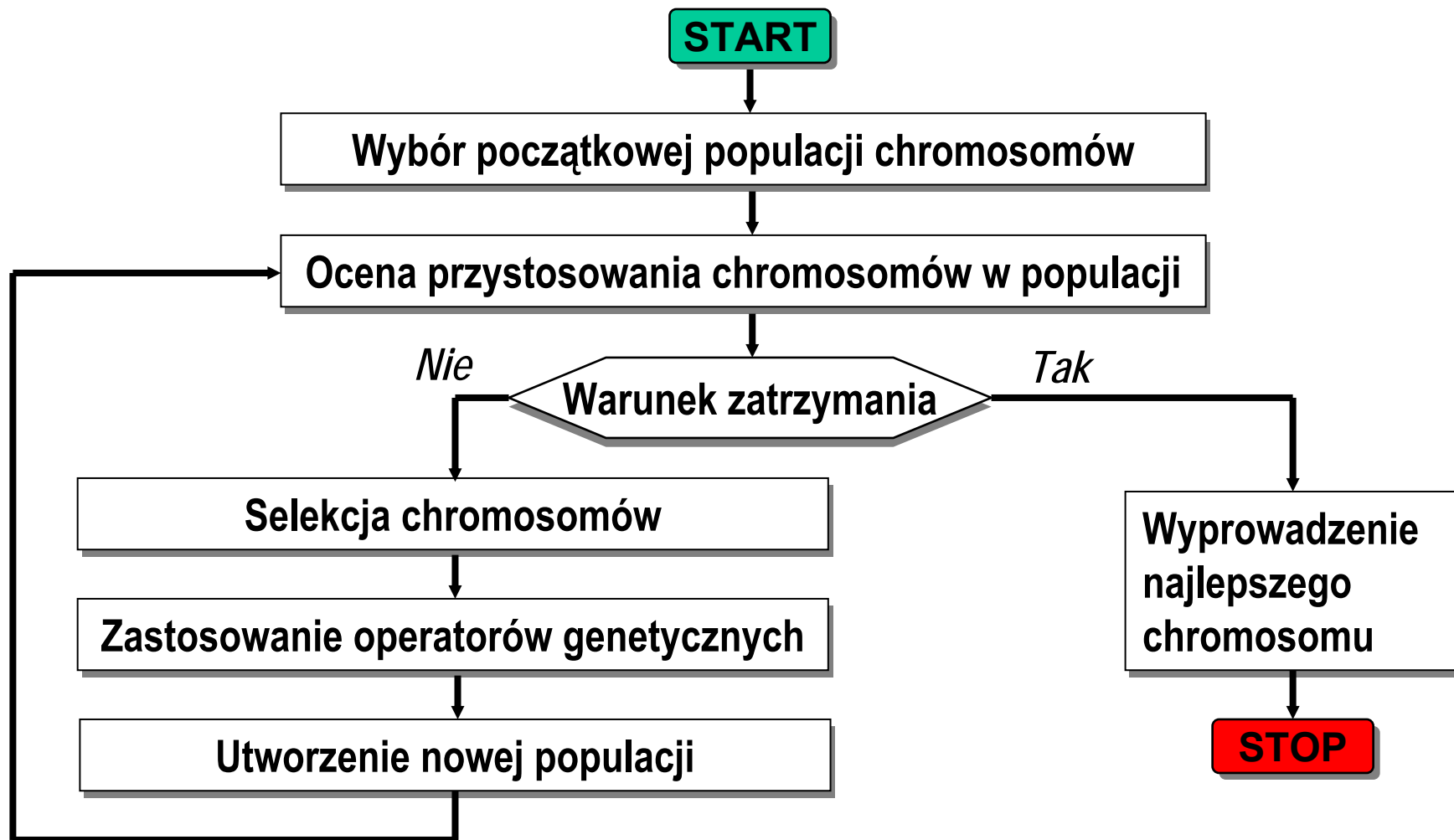
.....

- Korzystają tylko z funkcji celu, a nie z jej pochodnych (jak to ma miejsce w najczęściej stosowanych, gradientowych metodach optymalizacji).



Modele stosujące optymalizację genetyczną

Działanie algorytmów genetycznych





Modele stosujące optymalizację genetyczną

Działanie algorytmów genetycznych



- Wybór początkowej populacji chromosomów (zakodowanych wartości parametrów modelu) polega na losowym wyborze potrzebnej liczby chromosomów.
- Ocena przystosowania chromosomów w populacji polega na obliczeniu wartości funkcji celu dla każdego z tych chromosomów.
- Decyzja o zatrzymaniu obliczeń zależy od tego, czy spełniony jest narzucony przez użytkownika warunek, np. czy optymalizowana wartość dla najlepszego chromosomu nie zmienia się już znacząco.
- Selekcja chromosomów do następnej generacji odbywa się zgodnie z zasadą naturalnej selekcji, tzn. największe szanse na tworzenie nowego pokolenia chromosomów mają te z nich, które uzyskały najlepszą wartość funkcji celu. W wyniku procesu selekcji tworzy się populacja rodzicielska o liczebności równej poprzedniej populacji.



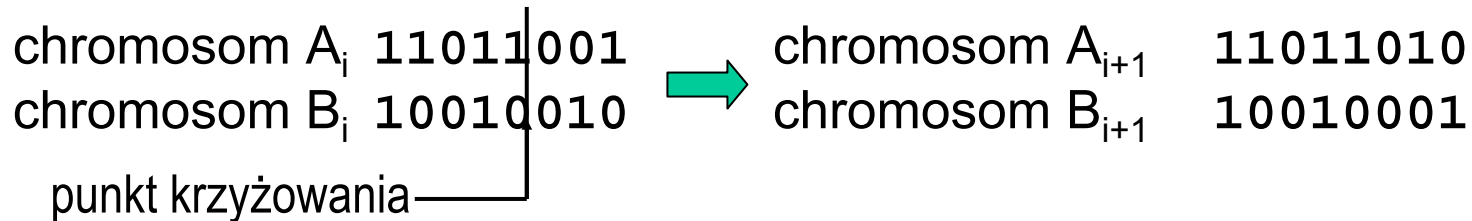
Modele stosujące optymalizację genetyczną

Działanie algorytmów genetycznych



Nowa populacja tworzona jest przez zastosowanie tzw. operatorów genetycznych

- *Krzyżowanie*. Polega na losowym kojarzeniu w pary chromosomów spośród wylosowanych do reprodukcji, wylosowaniu dla każdej pary pozycji w chromosomie określającej punkt krzyżowania, a następnie utworzeniu nowej pary z wymienionymi wzajemnie segmentami chromosomów. Np. dla pary chromosomów i -tego pokolenia:



- *Mutacja*. Polega na losowej zamianie pojedynczego genu (bitu) na przeciwny, np. 0 na 1. Liczba dokonywanych zmian jest niewielka, zwykle dotyczy kilku % wszystkich genów.



Modele stosujące optymalizację genetyczną

Obszary zastosowań w technice



- Harmonogramowanie produkcji
- Optymalizacja konstrukcji wyrobów
- Optymalizacja procesów produkcyjnych
- Optymalizacja parametrów eksploatacji urządzeń

Optymalizacja genetyczna ma bardzo szerokie zastosowania, m.in. z uwagi na brak ograniczeń postaci funkcji celu oraz możliwość znajdowania optimum globalnego (w przeciwieństwie do np. metod gradientowych prowadzących często do optimumów lokalnych, a więc rozwiązań gorszych).



Klasyfikacja bayesowska



Termin ten obejmuje systemy uczące się oparte na rachunku prawdopodobieństwa, wykorzystujące twierdzenie (wzór) Bayesa. Należą do nich:

- Klasyfikatory Bayesa
 - optymalny bez znaczenia praktycznego
 - naiwny klasyfikator Bayesa (NKB)
- Sieci bayesowskie



Klasyfikacja bayesowska

Twierdzenie Bayesa



Twierdzenie (wzór) Bayesa jest stosowane do zdarzeń, których zajście było zależne od czynników poprzedzających wystąpienie tych zdarzeń. Obliczane jest prawdopodobieństwo, na ile dany czynnik stanowi przyczynę wystąpienia zdarzenia (skutku).

$$P(h|D) = \frac{P(h) \cdot P(D|h)}{P(D)}$$

Wzór ten pozwala na obliczanie prawdopodobieństwa warunkowego słuszności hipotezy h dla danych D , jeżeli znane jest prawdopodobieństwo $P(D|h)$, tj. zaobserwowania danych D przy założeniu słuszności hipotezy h oraz bezwzględne prawdopodobieństwo zaistnienia tej hipotezy $P(h)$.



Naiwny klasyfikator Bayesa

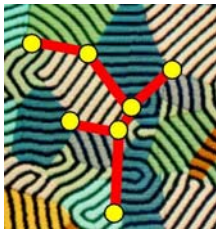
Zasady stosowania



Zastosowanie NKB wymaga wykonania obliczeń prawdopodobieństw na podstawie odpowiednio przygotowanego zbioru uczącego, składającego się z przykładów opisanych za pomocą atrybutów (wielkości wejściowych) i pojęcia docelowego (wielkości wyjściowej).

Działania wymagane dla zastosowania NKB:

- Ustalenie kategorii dla wielkości wejściowych oraz kategorii wielkości wyjściowej. NKB wymaga stosowania wielkości o charakterze nominalnym lub porządkowym, a nie ciągłym. „Wartość” atrybutu lub pojęcia docelowego jest określana przez oznaczenie jego przynależności do danej kategorii.



Naiwny klasyfikator Bayesa

Zasady stosowania



Działania wymagane dla zastosowania NKB c.d.:

- *Utworzenie zbioru uczącego*, składającego się z zapisów (rekordów) zawierających wartości atrybutów wszystkich wielkości wejściowych oraz odpowiadającej im wartości kategorii wielkości wyjściowej.
- Na podstawie zbioru uczącego *szacuje się prawdopodobieństwa* poszczególnych kategorii wielkości wyjściowej oraz prawdopodobieństwa poszczególnych wartości atrybutów (kategorii) wszystkich wielkości wejściowych dla poszczególnych kategorii wyjścia. Etap ten jest określany jako *uczenie* (trenowanie) klasyfikatora.



Naiwny klasyfikator Bayesa

Zasady stosowania



Działania wymagane dla zastosowania NKB c.d.:

- Na podstawie prawdopodobieństw oszacowanych na etapie uczenia, można obliczyć prawdopodobieństwa wystąpienia każdej z kategorii wielkości wyjściowej dla dowolnego przypadku (zestawu wielkości wejściowych).
- *Odpowiedzią* NKB jest ta wartość kategorii wielkości wyjściowej, która uzyskała największą wartość prawdopodobieństwa.



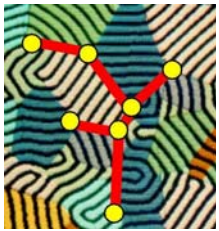
Naiwny klasyfikator Bayesa

Zasady stosowania



Obliczanie prawdopodobieństw w NKB:

- Zastosowanie klasyfikatora nie wymaga wyznaczenia prawdopodobieństwa $P(D)$ występującego w mianowniku wzoru Bayesa, gdyż wartość ta jest identyczna dla wszystkich możliwych kategorii wyjścia.
- Prawdopodobieństwa warunkowe $P(D|h)$ oraz $P(h)$ szacuje się na podstawie odpowiednich częstości występowania przykładów w zbiorze uczącym.
- Wzory służące do ich wyznaczenia są tak skonstruowane, aby uwzględnić sytuacje, w których niektóre kategorie nie są reprezentowane lub są bardzo nieliczne (automatyczne tworzenie 'wirtualnych' przypadków).



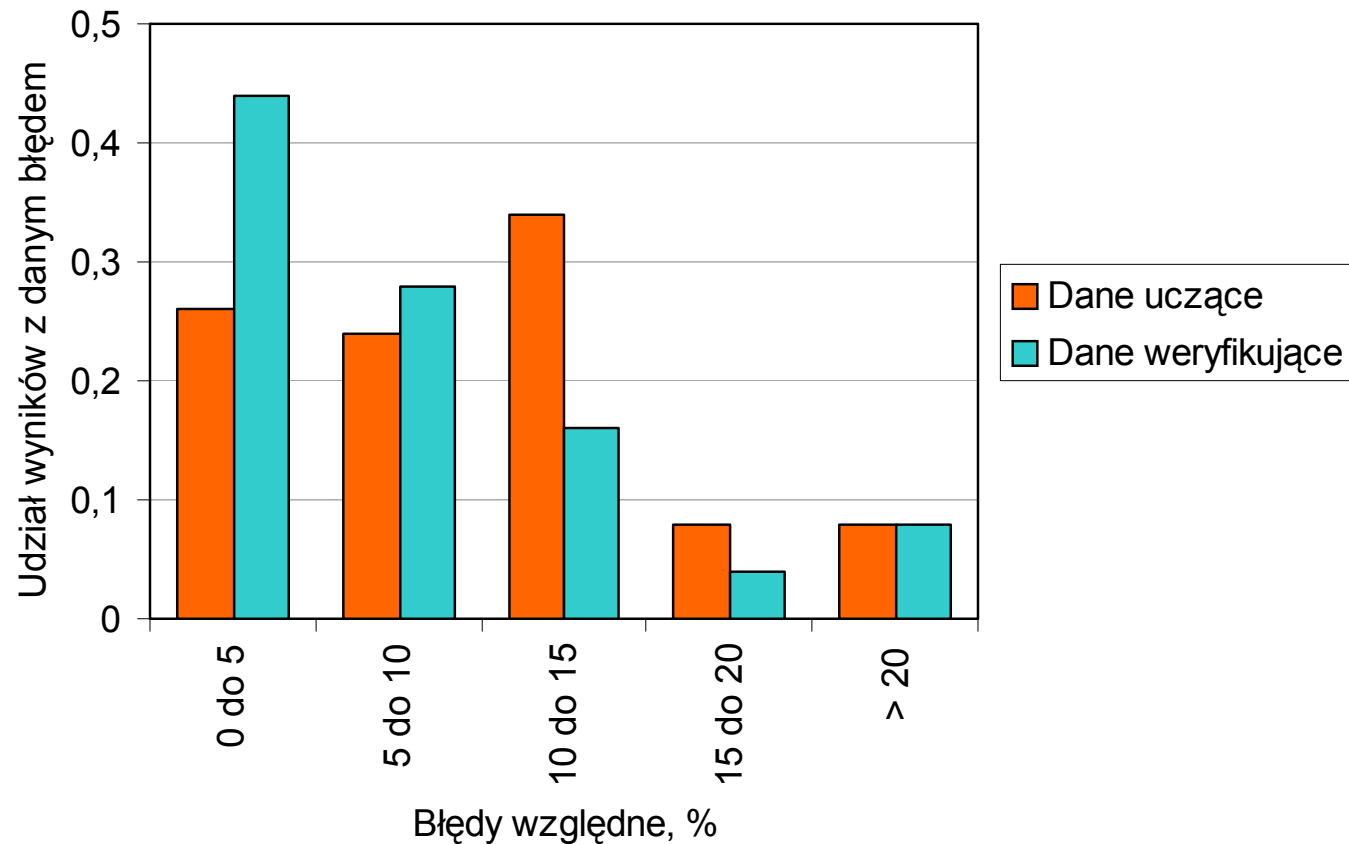
Naiwny klasyfikator Bayesa

Przykład zastosowania 1



Modelowanie
wytrzymałości żeliwa
sferoidalnego w
zależności od jego
składu chemicznego.

Zbiór uczący 700
rekordów,
weryfikujący
(niezależny) 90
rekordów





Naiwny klasyfikator Bayesa

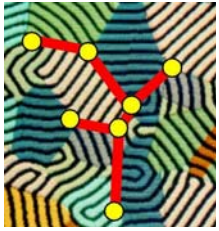
Przykład zastosowania 2



- Wielkościami wejściowymi były parametry procesu produkcyjnego związane z formą piaskową (12 parametrów), zaś wielkością wyjściową – wystąpienie wady typu porowatość gazowa w odlewie staliwnym.
- Dane uzyskano w jednej z odlewni krajowych produkującej odlewy w krótkich seriach, w formach z mas glinowych wilgotnych.
- W tym przypadku oryginalne wyjście było typu nieciągłego (kategoria „1” – brak wady, kategoria „2” – wystąpienie wady). Zbiór uczący zawierał 172 rekordy, zbioru weryfikującego nie było.

Wynik modelowania:

klasyfikator przewiduje poprawnie wystąpienie wady
w ok. 87% przypadków



MARSplines

Wielowymiarowe adaptacyjne regresyjne funkcje składowe

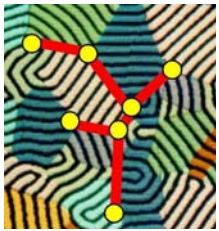


Nazwa MARSplines jest skrótem od angielskiego określenia Multivariate Adaptive Regression Splines (wielowymiarowe adaptacyjne regresyjne funkcje składowe).

Modele tego typu zostały opracowane przez J. H. Friedmana i zastosowane stosunkowo niedawno (1991–2001).

Realizują one zarówno zadania typu regresji, jak i klasyfikacji.

Stanowią bardzo dobre narzędzie w eksploracji danych, w szczególności w rozwiązywaniu trudnych problemów, tzn. gdy zależności między zmiennymi mają złożony charakter.



MARSplines

Charakterystyka ogólna



Modele MARSplines stosują nieparametryczną procedurę, w której nie jest zakładana żadna konkretna postać funkcyjnej zależności pomiędzy zmiennymi niezależnymi (wejściowymi), a zależnymi (wyjściowymi).

Zależność ta zbudowana jest ze zbioru współczynników i tzw. funkcji bazowych, które w całości są wyprowadzane z danych doświadczalnych.

Ogólna strategia tego modelu polega na podziale przestrzeni wejściowej na regiony, z których każdy posiada swoje własne równanie regresyjne (lub klasyfikacyjne).



MARSplines

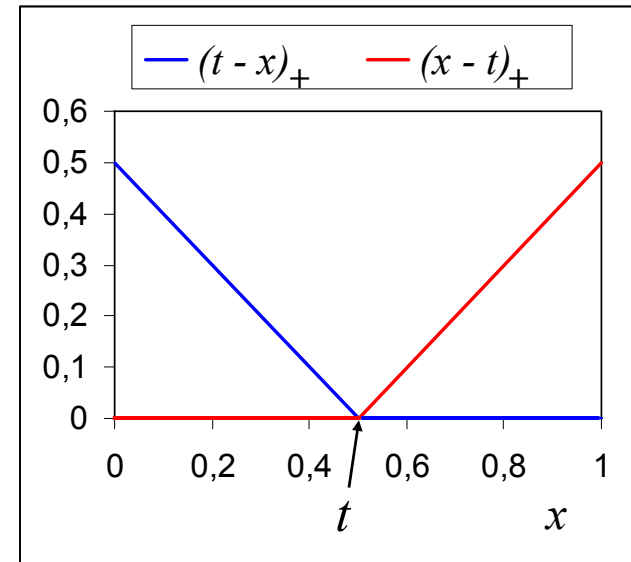
Podstawowe zależności modelu



Funkcje bazowe określane są jako dwustronnie obcięte funkcje dane zależnościami:

$$(x - t)_+ = \begin{cases} x - t & x > t \\ 0 & x \leq t \end{cases}$$

$$(t - x)_+ = \begin{cases} t - x & x < t \\ 0 & x \geq t \end{cases}$$



Parametr t jest tzw. węzłem funkcji bazowych (określającym „kawałki” regresji (kawałkami liniowej). Są one także wyznaczane z danych doświadczalnych.



MARSplines

Działanie procedury definiującej model



Algorytm modelu MARSplines przeszukuje przestrzeń wszystkich wielkości wejściowych. W trakcie tego działania do modelu dodawana jest coraz większa liczba funkcji bazowych tak, aby maksymalizować kryterium dopasowania (zminimalizować błąd średniokwadratowy) dla wszystkich danych.

Następnie algorytm stosuje technikę tzw. przycinania, która zmniejsza liczbę zastosowanych funkcji bazowych. W modelu pozostają jedynie te funkcje, które wnoszą zauważalny wkład w przewidywanie modelu. Przycinanie obejmuje także eliminację nieznaczących zmiennych wejściowych.



MARSplines

Najważniejsze cechy modelu



- Wysoka elastyczność powiązana z unikaniem nadmiernego dopasowania się do danych doświadczalnych dzięki zastosowanej technice przycinania
 - Uwaga: zbyt dokładne odwzorowanie danych uczących zmniejsza zdolność do uogólniania modelu, a więc przewidywania dla innych danych
- Nie jest wymagana żadna wiedza odnośnie postaci zależności między zmiennymi.
- Algorytm modelu automatycznie określa najważniejsze zmienne wejściowe oraz najbardziej znaczące interakcje między tymi zmiennymi.
- Istnieje możliwość jednoczesnego modelowania zależności kilku zmiennych wyjściowych w jednym modelu, wykorzystujących te same funkcje bazowe (analogia do sztucznych sieci neuronowych).



MARSplines

Możliwości zastosowań



Wymienione cechy modeli MARSpines stwarzają możliwości ich stosowania zbliżone do sztucznych sieci neuronowych oraz drzew decyzyjnych i wydają się być bardzo konkurencyjne dla obu tych modeli.

Nie są dotychczas znane aplikacje MARSpines do symulacji procesów metalurgicznych, jednak pojawiają się doniesienia o zastosowaniach w innych dziedzinach techniki, w tym do diagnostyki przyczyn zakłóceń procesów produkcyjnych.



METRO

MEtalurgiczny TRening *On-line*

Komputerowe metody analizy i sterowania procesami produkcyjnymi

Koniec wykładu



Edukacja i Kultura