

Artificial Intelligence, Artificial Life

(ALAI @ Cognitive Science.AMU)

© Maciej Komosinski

Institute of Computing Science
Poznan University of Technology

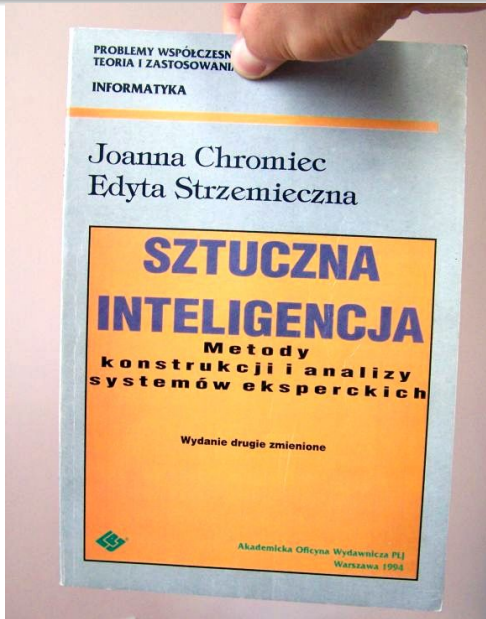
- Czym jest SI?
- Przeszukiwanie i optymalizacja
- Wnioskowanie
- Gry
- Odkrywanie wiedzy, uczenie maszynowe
 - k-NN
 - Drzewa i reguły decyzyjne
 - Sieci neuronowe
- Czym jest SŻ?
 - Modele życia, symulacja, ewolucja, teoria gier

- Nieobecni: systemy ekspertowe (SE), wspomaganie decyzji (WD), uczenie ze wzmocnieniem, przetwarzanie języka naturalnego, budowanie sztucznych mózgów

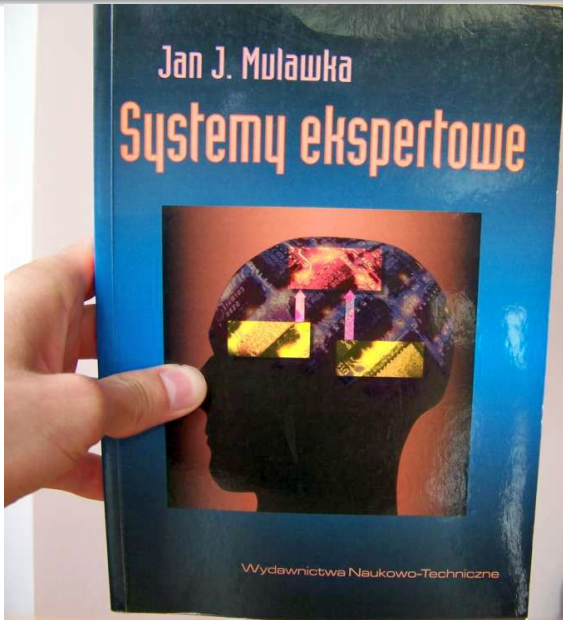
Co będzie przydatne?

- funkcja liniowa na wszystkie sposoby
- wielowymiarowe odległości, $|a - b|$
 - miara euklidesowa
 - miara Manhattan – „taksówkowa”
- logarytmy – przypadki szczególne (np. $\log_3 9$)
- używanie \sum i \prod
- pochodne – znaczenie
- teścik: zadania.pdf
- podstawy logiki i teorii prawdopodobieństwa: zdarzenia (nie)zależne, prawdopodobieństwo warunkowe, itp.

(1)



(2)



(3)



(4)



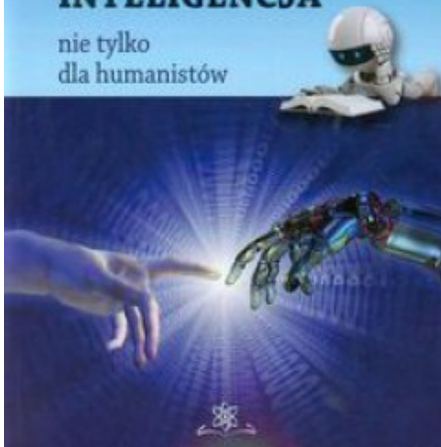
Dla chętnych zainteresowanych tematyką modelowania i symulacji życia, rozważających w przyszłości prowadzenie badań i rozwój naukowy – możliwy jest udział w spotkaniach grupy:

<http://www.cs.put.poznan.pl/mkomosinski/site/?q=artificial-life-laboratory>

Krzysztof FICOŃ

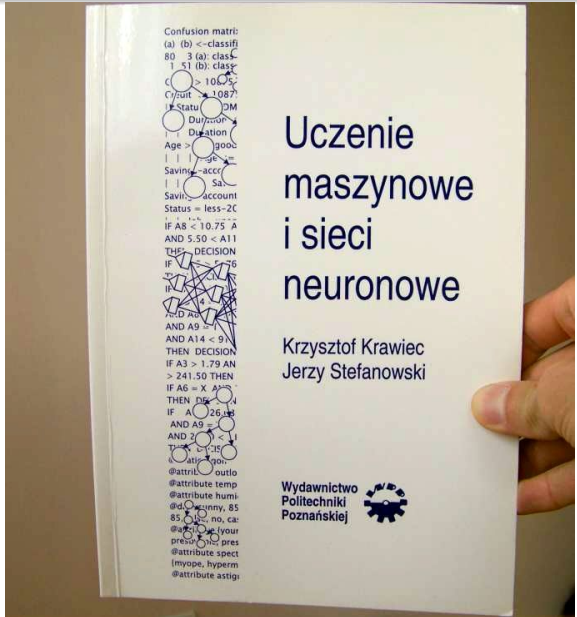
SZTUCZNA INTELIGENCJA

nie tylko
dla humanistów



(5)

(6)



(7)

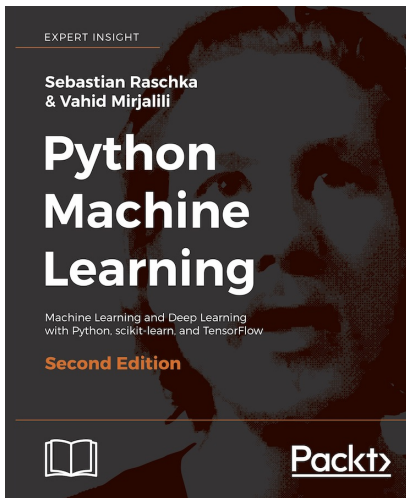
- Fragment materiałów do wykładu (Informatyka PP) „Obliczenia i Systemy Inspirowane Biologicznie”
- Sieci Neuronowe:
 - Skrypt R. Tadeusiewicza
winntbg.bg.agh.edu.pl/skrypty/0001/
 - Tutorial
www.k0pper.republika.pl/sieci.htm
 - Opis: działanie, uczenie, reguły
<http://www.poltynk.pl/marcin/dzialanie.html>
 - Podręcznik programu Statistica NN
<http://www.statsoft.pl/textbook/stneunet.html>

„Autostopem przez gALAIktykę”

(8) [http://www.cs.put.poznan.pl/ibladek/students/
alai/guide_opt.pdf](http://www.cs.put.poznan.pl/ibladek/students/alai/guide_opt.pdf)

[http://www.cs.put.poznan.pl/ibladek/students/
alai/guide_ml.pdf](http://www.cs.put.poznan.pl/ibladek/students/alai/guide_ml.pdf)

(9)



(10)



(11)



www.cs.put.poznan.pl/mkomosinski/materialy/

(12)

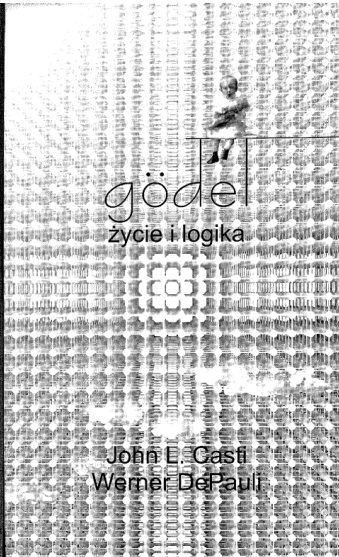
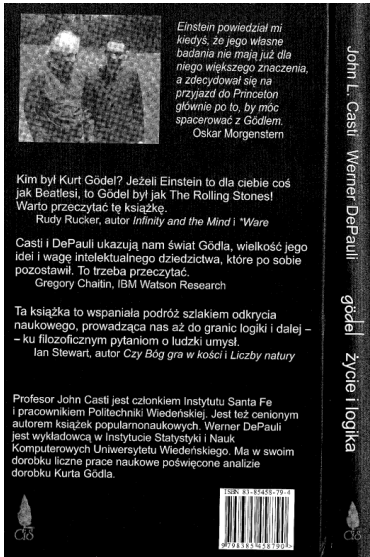
[https://www.youtube.com/channel/
UCCpnb4wSHw9pzxy5xNkGDdA/playlists](https://www.youtube.com/channel/UCCpnb4wSHw9pzxy5xNkGDdA/playlists)

maciej.komosinski@cs.put...

(13)

ważniak i własne notatki z zajęć.

Do poczytania:



Do obejrzenia:



Umiejętności po kursie ALAI

Lab i wykład

opanowanie pamięciowe < zrozumienie < umiejętność zastosowania

Jak się uczyć przedmiotów „ściśłych”?

Różnica między „prawie rozumiem”, „trochę rozumiem”, „chyba rozumiem” a „na pewno dobrze rozumiem” to różnica między dwójką a piątką!

Przygotowanie do egzaminu

- zrozumienie poruszanych zagadnień
 - pytania i wyjaśnianie wątpliwości na wykładzie i lab.
 - własne notatki z wykładu i lab. („co było”)
- opracowanie przykładowych *Zagadnień* ze [strony przedmiotu](#)
 - samodzielne rozwiązywanie zadań
 - pewność, że rozwiązania są dobre – weryfikacja
 - ew. opracowywanie pytań/zadań z *Zagadnień* po każdym wykładzie

- 1 What is Artificial Intelligence? [\(link\)](#) [waz1](#)
- 2 Optimization, complexity, MP [\(link\)](#) [pdf: OptIntroduction, MP](#)
- 3 BB, LS [pdf: BB-DP \(w/o DP\), LS](#)
- 4 SA, TS [pdf: SA, TS](#)
- 5 EA; multiple criteria [pdf: MetaheuristicsSummary](#)
- 6 Inductive reasoning, ML/KD/DM [waz9](#)
- 7 Unsupervised learning, k-means/k-medoids
- 8 Supervised learning, k-NN, Naive Bayes [waz10](#)
- 9 Decision trees and decision rules [waz10](#)
- 10 Neural networks, BP [waz12](#)
- 11 FL, inference, trees, games [waz2-6;8](#)
- 12 Evolution and co-evolution, advanced EA [\(link\)](#) [OIB_ALAI.pdf](#)
- 13 What is Artificial Life? Modeling life [OIB_ALAI.pdf](#)
- 14 Game theory [OIB_ALAI.pdf](#)

Główne zagadnienia poruszane na zajęciach

- 2-5. Klasy złożoności problemów optymalizacyjnych, ocena ich trudności, P i NP. Podejścia do znajdowania najlepszego rozwiązania; „krajobraz przystosowania”. Programowanie matematyczne. Algorytm BB i idea ograniczeń. Algorytmy *greedy* i *steepest*, sąsiedztwo – jego rozmiar i cechy. Algorytm SA i wpływ „temperatury”. Algorytm TS i działanie listy tabu. Algorytmy ewolucyjne, krzyżowanie, mutacja, selekcja. Algorytmy hybrydowe. Optymalizacja wielokryterialna, dominacja i pareto-optymalność rozwiązań, własności sumy ważonej.
6. Rodzaje uczenia, atrybuty i skale, tabela decyzyjna, regresja, klasyfikacja. Redukcja wymiarów; selekcja i konstrukcja atrybutów. Kryteria podziału algorytmów uczenia nadzorowanego. Reprezentacje wiedzy. Niedoskonałe dane. Walidacja wiedzy, miary jej oceny, macierz pomyłek, przeuczenie, „stopnie swobody”.
7. Grupowanie, k-means/k-medoids.
9. Miary informacji, obliczanie entropii, algorytm ID3, przycinanie (C4), generowanie reguł, drzewa a reguły, rozwój algorytmów DT.
11. Wiedza niedoskonała, zbiór rozmyty, logika rozmyta, FLC, rozmywanie, wyostrzanie, s-normy, p-normy. System wnioskowania, aksjomaty i reguły, podstawienie i unifikacja, predykaty. Wnioskowanie wstecz i w przód. Strategie ślepe (w głąb, wszerek) i poinformowane – najpierw-najlepszy (równomiernego kosztu (g), A^* ($g + h$) i zachłanna (h)). Zasada [minimaks](#), [alfa-beta](#) i MCTS.

- Does solving these problems require intelligence?
 - Timetabling, scheduling
 - Playing chess
 - Proving theorems
 - Autonomous mobile robot control
 - Medical diagnostics
 - Planning
- What comes to mind when we speak of AI?
Deep Blue (1996-7, IBM); DARPA Grand Challenge (2005); Watson (2011, IBM);
DRL for video games (2013, DeepMind); AlphaGo/Zero (2015+, DeepMind/Google);
Autonomous cars (2015+); Tay (2016, Microsoft)
- What is intelligence?
knowledge? inference? perception? cognition? learning? generalization? adaptation? creativity?
emotions? intuition? neurons? embodiment?
- How to define AI?
- Strong and weak AI
- What tasks is a computer/a human good/bad at?
- General and narrow AI
- Expert systems
- Examples in the first bullet point are tasks for AI: searching, inference, learning
- The search!
- Complexity of problems



Approaches to learning

- by heart
 - from instructions
 - deductive
 - by analogy
 - inductive
 - ...
-
- ML/KD/DM – induction.
 - supervised learning (“with a teacher”)
 - unsupervised learning (“without a teacher”)

Methodology of supervised learning

- The learning algorithm gets a set of pairs (x_i, y_i) , i.e., cases x_i assigned to some decision class y_i . The goal is to find the function f that correctly assigns cases to classes, $\forall i \ f(x_i) = y_i$.
- Goal: description, prediction
- Classification vs. regression
- Selection of attributes and construction of new ones (including PCA/SVD)
Exercises: <http://www.cs.put.poznan.pl/mkomosinski/site/?q=ml-attribute-space>
- The complexity of induced knowledge
- Underfitting and overfitting; regularization (L1 and L2, data augmentation, dropout, early stopping)
- Estimation of classification accuracy: single-run holdout, k-fold cross-validation, leaving-one-out
- Attributes and **scales**: nominal, ordinal, interval, ratio. Discretization, one-hot encoding.

Cechy (podział) algorytmów uczenia nadzorowanego

- strategia sterowania (*bottom-up*, *top-down*)
- rodzaj opisu (charakteryzujący, dyskryminujący)
- sposób uczenia (inkrementacyjny, całościowy)
- reprezentacja danych wejściowych (tabela decyzyjna, szereg czasowy, relacja, gramatyka, ...) i wyjściowych (reguły decyzyjne, drzewa, grafy, ...)
- radzenie sobie z niedoskonałymi danymi (brakujące wartości, niespójności, klasy niezrównoważone) i różnymi skalami atrybutów

Unsupervised learning. Clustering with k-means

- Cel: podzielić n obserwacji $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$ na k grup (S_1, S_2, \dots, S_k)
- ...tak, żeby każda obserwacja należała do grupy z najbardziej podobną średnią, czyli żeby minimalne było

$$\sum_{i=1}^k \sum_{\mathbf{x}_j \in S_i} \|\mathbf{x}_j - \boldsymbol{\mu}_i\|^2$$

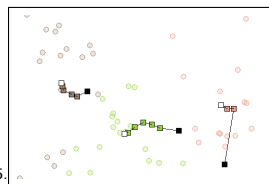
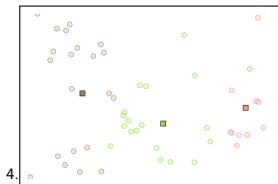
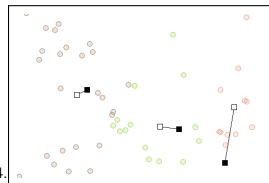
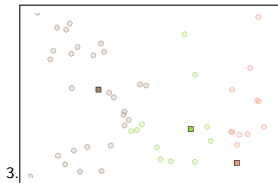
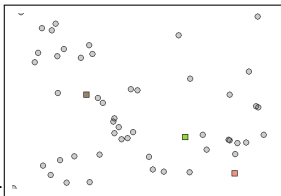
gdzie $\boldsymbol{\mu}_i$ jest średnią z grupy S_i .

- Niestety, znalezienie takiego podziału jest NP-trudne – dlatego stosuje się heurystyki.
- A jeśli nie wiemy, jakie „powinno” być k ?

[https://en.wikipedia.org/wiki/Silhouette_\(clustering\)](https://en.wikipedia.org/wiki/Silhouette_(clustering))

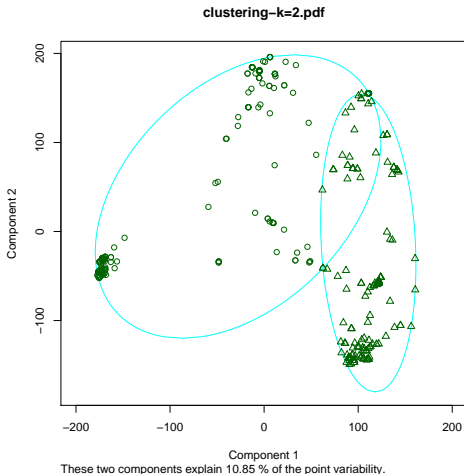
Unsupervised learning. Clustering with k-means

- 1 Wybieramy k (np. $k = 3$)
- 2 Inicjalizujemy losowo k centroidów
- 3 Przypisujemy dane do najbliższych centroidów
- 4 Przesuwamy centroidy do środka „swoich” danych
- 5 Powtarzamy 3 i 4 aż centroidy się ustabilizują



Unsupervised learning. Clustering with k-medoids

Example – 381 cases, hundreds of attributes, PCA 2D view:



Supervised learning

Honest



Tridas



Vickie



Mike

Crooked



Wally



Waldo



Barney

Prediction? Concept?



Tridas



Vickie



Mike

Honest =

k-NN

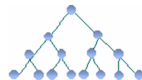
- Ciągłe liczy odległości...
 - Model zawiera całość danych
 - Jak dobrać k ?
 - IBL2 i kolejne (David Aha) tworzą mniejszy model
-
- Ćwiczenia: <http://www.cs.put.poznan.pl/mkomosinski/site/?q=ml-kNN-IBL>

Naive Bayes classifier

- Prawdopodobieństwo warunkowe
 - Twierdzenie Bayesa
 - „Naiwność”
 - Problem zerowych prawdopodobieństw
 - Dobry przykład policzony na *ważniaku*
-
- Ćwiczenia: <http://www.cs.put.poznan.pl/mkomosinski/site/?q=ml-bayes>

Decision trees: data

	<u>height</u>	<u>hair</u>	<u>eyes</u>	<u>attractiveness</u>
1	short	blond	blue	+
2	tall	blond	brown	-
3	tall	red	blue	+
4	short	dark	blue	-
5	tall	dark	blue	-
6	tall	blond	blue	+
7	tall	dark	brown	-
8	short	blond	brown	-



Entropia: obliczenia wstępne

- Znaczenie i jednostki entropii
- Entropia dla „sprawiedliwej” monety
- Entropia dla 8-ściennej sprawiedliwej kostki
- ...a gdyby jedna ścianka była nieco większa?
- Entropia dla n -ściennej sprawiedliwej kostki
- Entropia dla monety, dla której jedna strona nigdy nie wypadnie

Miary informacji

- Częściowa entropia warunkowa

$$H(Y | x_i) = - \sum_{j=1}^m p(y_j | x_i) \cdot \log p(y_j | x_i)$$

- Pełna entropia warunkowa

$$H(Y | X) = \sum_i p(x_i) \cdot H(Y | x_i)$$

- Miary zysku na informacji dla atrybutu i -tego

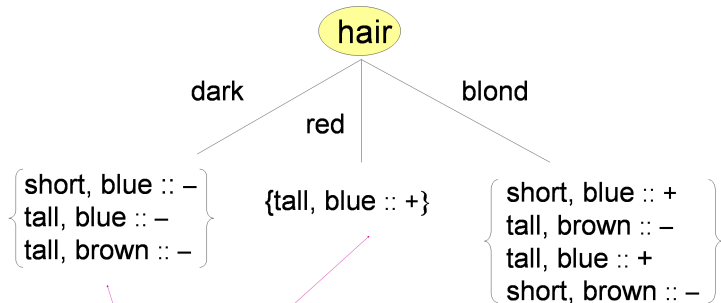
$$\text{Info.gain}(a_i) = H(d) - H(d | a_i)$$

$$\text{Info.gain.ratio}(a_i) = \frac{\text{Info.gain}(a_i)}{H(a_i)}$$

Obliczenia – przykład „Attractiveness”

- $H(\text{Attractiveness}) =$
 - $H(\text{Attractiveness}|\text{Height}) =$
 - $H(\text{Attractiveness}|\text{Hair}) =$
 - $H(\text{Attractiveness}|\text{Eyes}) =$
 - ... i tak do wyczerpania możliwości porządkowania przypadków.
 - Dlaczego warto używać Gain.ratio?
 - Przycinanie w C4: $\frac{E}{N} + \alpha L$
-
- Ćwiczenia: <http://www.cs.put.poznan.pl/mkomosinski/site/?q=ml-decision-tree-induction>

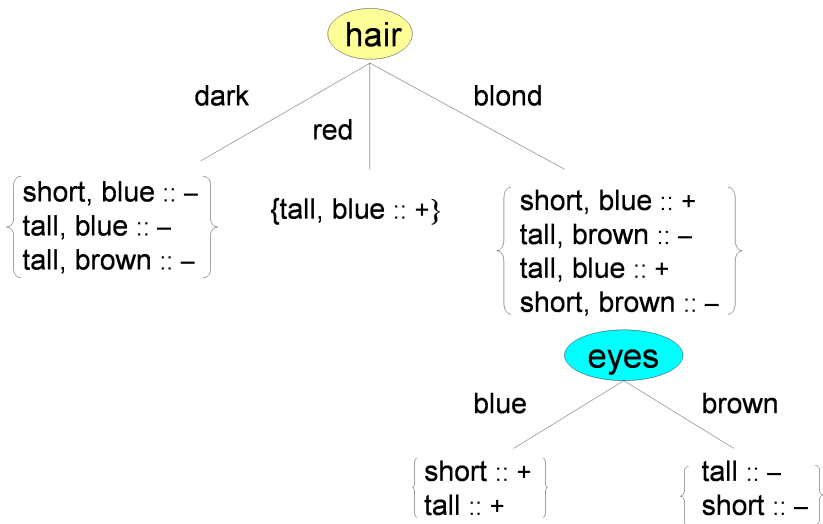
Building a decision tree (1)



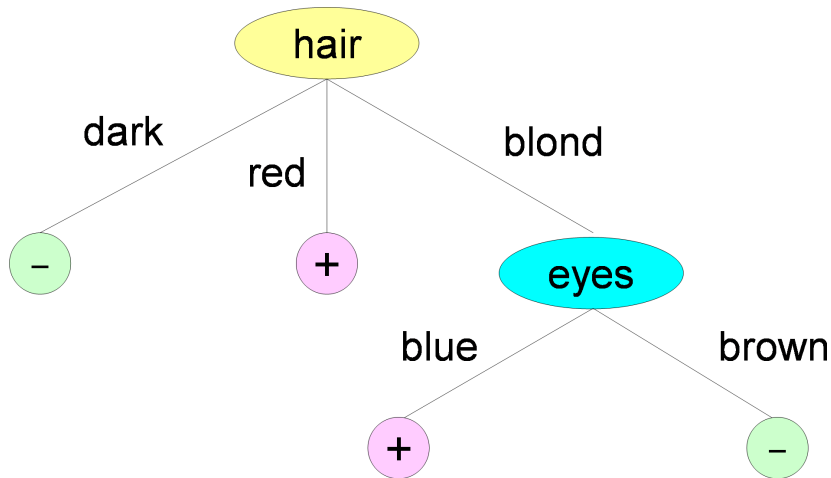
Completely classifies dark-haired
and red-haired people

Does not completely classify
blond-haired people.
More work is required

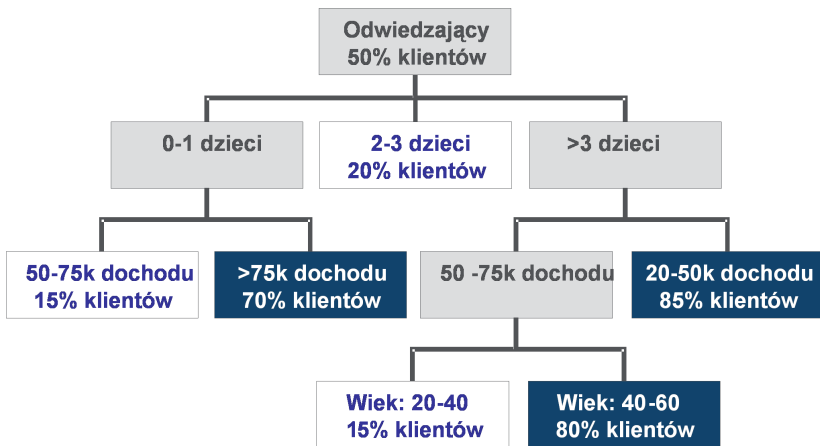
Building a decision tree (2)



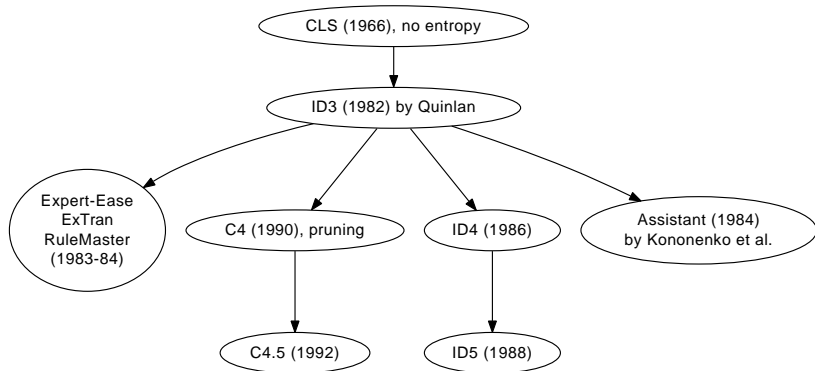
Decision tree: learned predictive model



Drzewa decyzyjne – przykład „socjologiczny”



Drzewa decyzyjne – rozwój algorytmów indukcji

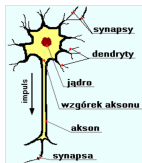


Drzewa decyzyjne a reguły decyzyjne: które może więcej?

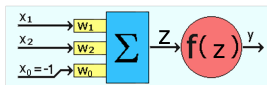
- drzewo \rightarrow zestaw reguł ?
- zestaw reguł \rightarrow drzewo ?

Sieci neuronowe – model

neuron biologiczny



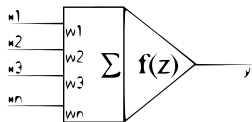
model



- dendryty \Leftrightarrow wejścia $\mathbf{x} = [x_0, x_1, x_2, \dots, x_n]^T$
 - synapsy \Leftrightarrow wagi $\mathbf{w} = [w_0, w_1, w_2, \dots, w_n]^T$
 - błona komórkowa \Leftrightarrow sumator Σ
 - wzgórek aksonu \Leftrightarrow blok aktywacji $f(z)$
 - akson \Leftrightarrow wyjście y
-
- łączne pobudzenie $z = \mathbf{w}^T \mathbf{x}$
 - funkcja aktywacji $y = f(z)$

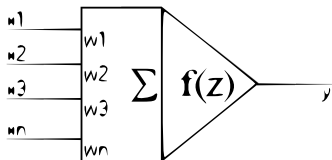
Sieci neuronowe – oznaczenia

- $x_0, x_1, x_2, x_3, \dots$ – wartości sygnałów wejściowych (x_0 podłączony do stałej)
- $w_0, w_1, w_2, w_3, \dots$ – wagi przypisane wejściom
- z – PSP: suma ważona sygnałów wejściowych ($z = x_0 \cdot w_0 + x_1 \cdot w_1 + \dots + x_n \cdot w_n$)
- $f()$ – funkcja aktywacji
- y – wyjście neuronu, $y = f(z)$.
Pochodna $y = \frac{1}{1+e^{-z}}$ wynosi $y \cdot (1 - y)$
- o – oczekiwana wartość wyjścia neuronu
- E – błąd dla pojedynczego wzorca, np.
$$E = \frac{(o-y)^2}{2}$$



Sieci neuronowe – budowa i działanie

- architektury SN (jednokierunkowa, rekurencyjna, Hopfielda, ...)
- analiza dwu-wejściowego neuronu progującego; rola w_0
- funkcje aktywacji (liniowa, progowa, sigmoidalna, (leaky) ReLU, i inne)
- atrybuty nominalne
- więcej wejść... więcej wyjść... więcej warstw
- przeuczenie
- typowe zastosowania
- nowoczesne architektury: konwolucyjne SN (video),
głębokie SN (video)



Sieci neuronowe – uczenie (algorytm gradientowy) (ulepszenia)

$$w_{nowa} = w_{stara} + \Delta w$$

$$\Delta w = -\eta \frac{\partial E}{\partial w}$$

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial z} \frac{\partial z}{\partial w}$$

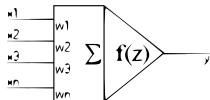
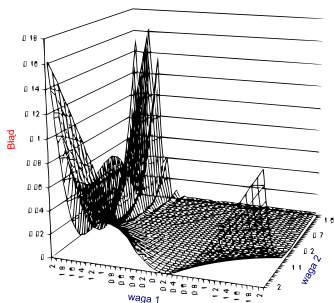
$$\delta = -\frac{\partial E}{\partial z} \text{ (tzw. „sygnał błędny”)}$$

$$\frac{\partial z}{\partial w} = x$$

$$\Delta w = \eta \delta x$$

$$\delta = -\frac{\partial E}{\partial z} = -\frac{1}{2} \cdot \frac{\partial}{\partial z} (o - y)^2 = (o - y) \frac{\partial y}{\partial z} = (o - y) f'(z)$$

$$\Delta w = \eta \cdot (o - y) f'(z) \cdot x$$



Sieci neuronowe – wsteczna propagacja błędu (1969, 1974)

- Wiele warstw – algorytm *back propagation* (BP, BPROP)
- Dla warstwy wyjściowej: $\delta = (o - y) \cdot f'(z)$
- Dla pozostałych: delta jest równa sumie ważonej delt z warstwy następnej przemnożonej przez pochodną funkcji aktywacji:

$$\delta_j = f'(z) \cdot \sum_{k=1}^K \delta_{(j+1)k} w_{(j+1)k}$$

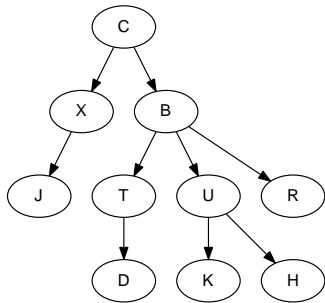
j – aktualnie analizowana warstwa neuronów

$(j + 1)$ – następna warstwa neuronów

K – liczba neuronów w warstwie $(j + 1)$

k – neuron z warstwy $(j + 1)$

Przeszukiwanie drzewa wszerz i w głąb – przykład




wszerz (FIFO)

węzeł	kolejka
	C
C	XB
X	BJ
B	JTUR
J	TUR
T	URD
U	RDKH
R	DKH
D	KH
K	H
H	(pusta)


w głąb (LIFO)

węzeł	kolejka
	C
C	XB
X	JB
J	B
B	TUR
T	DUR
D	UR
U	KHR
K	HR
H	R
R	(pusta)

Architektury ewolucyjne a optymalizacja

- ewolucja i jej mechanizmy 
- lamarkizm, darwinizm, efekt Baldwina, gradualizm, saltacjonizm, punktualizm
- ewolucja połączona z uczeniem; algorytmy hybrydowe (memetyczne, GLS)
- ewolucja z funkcją celu wewnętrzną / zewnętrzną; kierowana przez człowieka; ograniczona / otwarta
- koewolucja
 - konkurencyjna / kooperatywna
 - efekt Czerwonej Królowej, wyścig zbrojeń, kiepskie stany stabilne (MSS)
 - patologie: brak lub utrata gradientu, cykle
 - przeciwdziałanie: CFS, HOF

Sztuczne Życie

- Czym jest?
- Definicje życia 
- Obszary badań
- Przykładowe modele
 - Automaty komórkowe (CA)
 - Systemy wieloagentowe (MAS)
 - L-systemy
 - Symulacje VR
 - → Robotyka: poziomy autonomii robotów
- Pojęcia: *CAS, emergence, stigmergy, situatedness, embodiment, sensory-motor coordination, perceptual aliasing, active perception*
- Symulatory SŻ: Starlogo i Netlogo, MCell, Framsticks, Avida, Repast, ...
- Proste symulacje ilustrujące: www.alife.pl
- Artykuły: www.alife.pl/czym-jest-sztuczne-zycie ;
www.alife.pl/zycie-w-komputerze-symulacja-czy-rzeczywistosc ;
en.wikipedia.org/wiki/Simulated_reality