

# Uczenie maszynowe: *wykład 1*

Paweł Cichosz

1 Informacje organizacyjne

2 Tematyka przedmiotu

3 Indukcyjne uczenie się

# Informacje o przedmiocie

Strona przedmiotu: <http://elektron.elka.pw.edu.pl/~pcichosz/uma>

## Literatura:

- 1 T. Mitchell (1997): *Machine Learning*.
- 2 I. H. Witten, E. Frank, M. A. Hall (2016): *Data Mining: Practical Machine Learning Tools and Techniques* (4th edition).
- 3 Y. S. Abu-Mostafa, Y.S., M. Magdon-Ismael, H.-T. Lin (2012): *Learning from Data: A Short Course*.
- 4 M. Muraszewicz, R. Nowak, red. (2022): *Sztuczna inteligencja dla inżynierów: metody ogólne*.
- 5 P. Cichosz (2000, 2007): *Systemy uczące się*.
- 6 P. Cichosz (2015): *Data Mining: Explained Using R*.

# Zasady zaliczania

**Kolokwia:**  $2 \times 25$  punktów

- w połowie semestru (termin wykładu 7 lub 8),
- na końcu semestru (termin wykładu 14 lub 15).

**Projekt:**  $10 + 40$  punktów

- 1 zespoły dwuosobowe,
- 2 implementacja i zbadanie właściwości algorytmu uczenia się,
- 3 ogłoszenie tematów w ciągu 2-3 tygodni,
- 4 dokumentacja wstępna do końca 7. tygodnia semestru,
- 5 zakończenie do końca przedostatniego tygodnia semestru.

# Formuła wykładu

Przeplatające się tryby:

Slajdy:

- bardziej uporządkowany wywód,
- czasem mniejsza wnikliwość,
- mniej interakcji ze słuchaczami,
- większa dyscyplina czasowa.

Kreda/marker i tablica:

- czasem większa wnikliwość,
- więcej interakcji ze słuchaczami,
- zwiększone ryzyko pomyłek,
- nie zawsze wszystko zmieści się w czasie.

W obu przypadkach:

- dostępne slajdy lub notatki podsumowujące treść wykładów,
- ograniczony zakres i spokojne tempo,
- główne treści ilustrowane przykładami.

- 1 Informacje organizacyjne
- 2 **Tematyka przedmiotu**
- 3 Indukcyjne uczenie się

## W stronę definicji uczenia się

Charakterystyka systemów uczących się: autonomiczna poprawa jakości wykonywania zadania na podstawie zewnętrznej informacji trenującej.

Program uczący się:

- algorytm wykonywania zadania (niedoskonały, niepełny),
- algorytm uczenia się (dostarczający wiedzy/umiejętności uzupełniającej/doskonalałej algorytm wykonywania zadania).

## W stronę definicji uczenia się

### Rodzaje informacji trenującej:

- przykłady etykietowane (uczenie się z nauczycielem / z nadzorem),
- odpowiedzi na zapytania (uczenie się z nauczycielem / z nadzorem),
- przykłady nieetykietowane (uczenie się bez nauczyciela / bez nadzoru),
- nagrody/kary (uczenie się ze wzmocnieniem).

### Wynik uczenia się:

- wiedza (jakie coś jest?),
- umiejętność (jak coś zrobić?).



# Zakres przedmiotu

- 1 Elementy teorii uczenia się:
  - 1 model PAC,
  - 2 spójne uczenie się,
  - 3 agnostyczne uczenie się,
  - 4 wymiar Vapnika-Chervonenkisa,
  - 5 przestrzeń wersji, generalizacja i specjalizacja.
- 2 Podstawowe algorytmy uczenia się klasyfikacji i regresji:
  - 1 indukcja reguł,
  - 2 indukcja drzew decyzyjnych,
  - 3 naiwny klasyfikator bayesowski,
  - 4 regresja i klasyfikacja liniowa.
- 3 Ocena jakości modeli klasyfikacji i regresji.
- 4 Wybrane bardziej zaawansowane algorytmy klasyfikacji i regresji:
  - 1 las losowy,
  - 2 SVM.
- 5 Uczenie się ze wzmocnieniem.

# Ograniczenia przedmiotu

## Koncentracja na podstawach:

- brak algorytmów bardziej zaawansowanych,
- niektóre algorytmy prezentowane bez wyprowadzenia lub uzasadnienia,
- niektóre algorytmy prezentowane ze względu na znaczenie historyczne lub walory dydaktyczne.

Brak sieci neuronowych: omawiane na innych przedmiotach.

Brak szerszej dyskusji zagadnień związanych z zastosowaniami: omawiane na innych przedmiotach.

- 1 Informacje organizacyjne
- 2 Tematyka przedmiotu
- 3 Indukcyjne uczenie się

# Podstawowa terminologia i notacja

**Dziedzina:**  $X$  – zbiór obiektów, do których ma się odnosić pozyskiwana wiedza.

**Przykład:**  $x \in X$  – dowolny element dziedziny.

**Atrybuty:**  $a_1, a_2, \dots, a_n$ ,  $a_i : X \rightarrow A_i$  – wektorowa reprezentacja przykładów:

**ciągłe (numeryczne):** reprezentacja i interpretacja liczbowa (możliwa arytmetyka),

**dyskretne (kategoryczne):** brak interpretacji liczbowej:

**nominalne:** bez relacji porządku (możliwe porównania  $=$ ,  $\neq$ ),

**porządkowe:** z relacją porządku (możliwe porównania  $=$ ,  $\neq$ ,  $<$ ,  $\leq$ ,  $>$ ,  $\geq$ ).

**Wektor wartości atrybutów:**

$a_1(x)$	$a_2(x)$	$\dots$	$a_n(x)$
----------	----------	---------	----------

## Zadanie klasyfikacji (uczenia się pojęć)

**Pojęcie:**  $c : X \rightarrow C$ ,  $C$  – skończony zbiór klas.

**Zbiór trenujący:**  $T \subseteq D \subset X$ , dla  $x \in D$  znane  $c(x)$ .

**Model:**  $h : X \rightarrow C$ ,  $h \approx c$ .

**Predykcja:** wyznaczenie  $h(x)$  dla  $x \in X$  – zastosowanie modelu.

**Klasa pojęć:**  $\mathbb{C}$  – zbiór pojęć rozważanych dla pewnej dziedziny (podzbiór wszystkich możliwych pojęć spełniający określone warunki, determinuje złożoność *zadania* uczenia się).

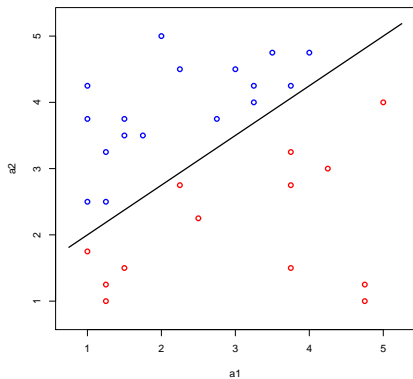
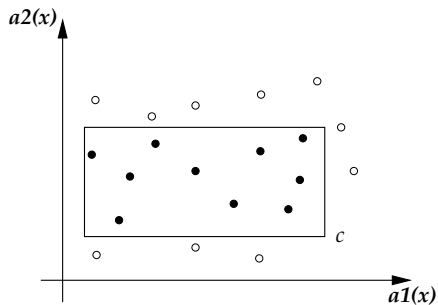
**Przestrzeń modeli:**  $\mathbb{H}$  – zbiór możliwych modeli rozważanych dla pewnej dziedziny (np. wyznaczony przez określenie reprezentacji), determinuje złożoność *algorytmu* uczenia się.

**Uczenie się jako przeszukiwanie:** znaleźć model  $h \in \mathbb{H}$  najlepiej przybliżający pojęcie  $c \in \mathbb{C}$ .

## Przykładowe dziedziny, klasy pojęć i przestrzenie modeli

- Prostokąty:** Dziedzina  $\mathcal{R}^2$  (punkty na płaszczyźnie), atrybuty – współrzędne kartezjańskie, pojęcia i modele – reprezentowane przez prostokąty o bokach równoległych do osi układu współrzędnych (przykłady pozytywne wewnątrz i na brzegu prostokąta, przykłady negatywne na zewnątrz prostokąta).
- Proste:** Dziedzina  $\mathcal{R}^2$  (punkty na płaszczyźnie), atrybuty – współrzędne kartezjańskie, pojęcia i modele – reprezentowane przez proste (przykłady pozytywne po dodatniej stronie prostej i na prostej, przykłady negatywne po ujemnej stronie prostej).
- Koniunkcje boolowskie:** Dziedzina  $\{0, 1\}^n$  (łańcuchy binarne), atrybuty – bity, pojęcia i modele – reprezentowane przez koniunkcje boolowskich literałów ( $a_i, \neg a_i$ ).

## Przykładowe dziedziny, klasy pojęć i przestrzenie modeli



# Zadanie regresji

Funkcja docelowa:  $f : X \rightarrow \mathcal{R}$ .

Zbiór trenujący:  $T \subseteq D \subset X$ , dla  $x \in D$  znane  $f(x)$ .

Model:  $h : X \rightarrow \mathcal{R}$ ,  $h \approx f$ .



## Jakość modelu

Błąd klasyfikacji na zbiorze:

$$e_{S,c}(h) = \frac{|\{x \in S \mid h(x) \neq c(x)\}|}{|S|}$$

Błąd rzeczywisty klasyfikacji:

$$e_{\Omega,c}(h) = P(h(x) \neq c(x) \mid x \in X, x \sim \Omega)$$

$\Omega$  – rozkład prawdopodobieństwa na dziedzinie  $X$ .

Błąd średniokwadratowy na zbiorze (MSE, *mean square error*):

$$\text{mse}_{S,f}(h) = \frac{\sum_{x \in S} (f(x) - h(x))^2}{|S|}$$

Błąd średniokwadratowy rzeczywisty:

$$\text{mse}_{\Omega,f}(h) = \mathbf{E}[(f(x) - h(x))^2 \mid x \in X, x \sim \Omega]$$

# Obciążenie indukcyjne

- Właściwości algorytmu uczenia się, które determinują wybór  $h$  na podstawie  $T$  (sposób uogólnienia przykładów trenujących).
- Konieczne w celu wyboru jednej z wielu możliwych generalizacji.
- Realizacja:
  - **obciążenie reprezentacji**: zawężenie przestrzeni modeli dostępnej dla algorytmu,
  - **obciążenie preferencji**: wprowadzenie kryteriów preferencji modeli (np. preferencja dla najprostszych modeli).

# Nadmierne dopasowanie

- Model  $h_1 \in \mathbb{H}$  jest nadmiernie dopasowany do  $T$ , jeśli istnieje model  $h_2 \in \mathbb{H}$  taki, że  $h_1$  ma mniejszy błąd na  $T$ , ale  $h_2$  ma mniejszy błąd rzeczywisty.
- Zasadnicze wyzwanie przy tworzeniu modeli przez uczenie się.
- Obliczeniowa teoria uczenia się charakteryzuje poziom ryzyka nadmiernego dopasowania w zależności od przestrzeni modeli i właściwości algorytmów.
- Zapobieganie i kontrola:
  - stosowanie algorytmów o mniejszym ryzyku nadmiernego dopasowania,
  - wprowadzanie do algorytmów mechanizmów zmniejszających ryzyko nadmiernego dopasowania,
  - staranna ocena jakości modeli.

## Przykład: koniunkcje boolowskie

$a_1$	$a_2$	$a_3$	$c$	$h_1$	$h_2$
<b>0</b>	<b>0</b>	<b>0</b>	0	0	0
0	0	1	1	0	0
<b>0</b>	<b>1</b>	<b>0</b>	0	0	0
0	1	1	1	0	1
1	0	0	0	1	0
<b>1</b>	<b>0</b>	<b>1</b>	1	1	0
1	1	0	0	1	0
1	1	1	1	1	1

- $n = 3, c = a_3$
- $T = \{000, 010, 101\}$
- $h_1 = a_1, h_2 = a_2 \wedge a_3$
- $\Omega$  – rozkład jednostajny na dziedzinie  $\{0, 1\}^3$
- $e_{T,c}(h_1) = 0, e_{T,c}(h_2) = \frac{1}{3}$
- $e_{\Omega,c}(h_1) = \frac{4}{8}, e_{\Omega,c}(h_2) = \frac{2}{8}$