

*SD*

Inżynieria Mechatroniczna

Podstawy Sztucznej Inteligencji i Ucznienia Głębokiego:  
**2: Regresja i Klasyfikacja**

Ziemowit Dworakowski  
*AGH w Krakowie*

1

---

---

---

---


---

---




---

---

**Przykład: Zakup używanego samochodu** *SD*

 Mamy 7-letniego Opla Astrę, z 80 000 km przebiegu. Chcemy sprzedać go szybko ale za jak najwyższą cenę

**Cena początkowa**

 ← Za niska  Za wysoka 

Niewiele zarobimy... Samochód nie sprzedaje się szybko

2

---

---

---

---

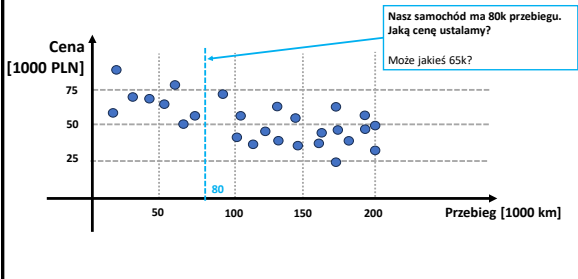
---

---

---

---

**Przykład: Zakup używanego samochodu** *SD*



Nasz samochód ma 80k przebiegu. Jaką cenę ustalamy?  
Może jakieś 65k?

3

---

---

---

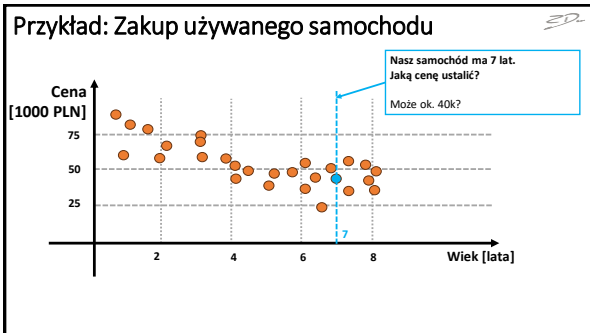
---

---

---

---

---



4

---

---

---

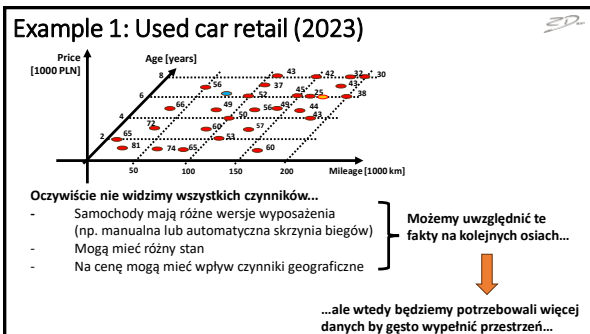
---

---

---

---

---



5

---

---

---

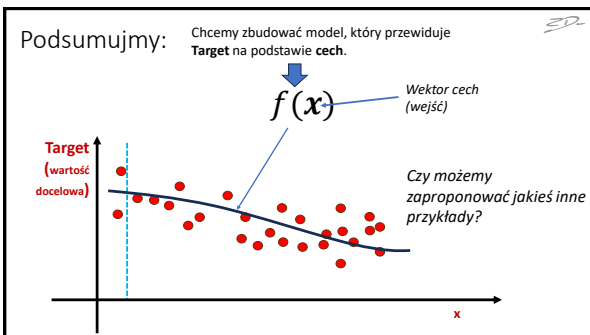
---

---

---

---

---



6

---

---

---

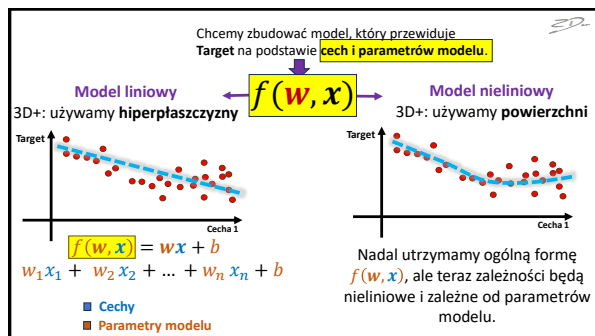
---

---

---

---

---



7

---

---

---

---

---

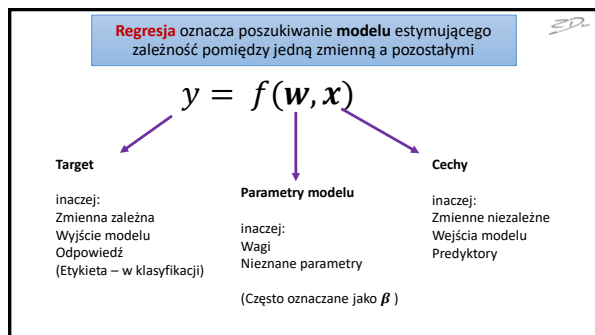
---

---

---

---

---



8

---

---

---

---

---

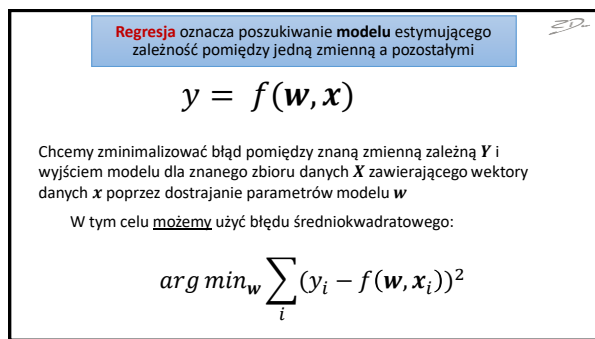
---

---

---

---

---



9

---

---

---

---

---

---

---

---

---

---

### Regresja liniowa

Chcemy znaleźć taką linię, że suma kwadratów zielonych odcinków jest jak najmniejsza

**Model:**  
 $y = w_1x + w_b$  →  $arg \min_w \sum_i (y_i - (w_1x_i + w_b))^2$

**Dopasowanie modelu:**

10

---

---

---

---

---

---

---

---

### Estymator lokalnie liniowy

Linia prosta często nie pozwala na dobre modelowanie danych ...

Ale to na czym nam naprawdę zależy, to skuteczność przewidywania wartości dla konkretnego  $x$

Więc może jesteśmy w stanie zachować model ale zmienić metodę dopasowania?

11

---

---

---

---

---

---

---

---

### Estymator lokalnie liniowy

Chcemy zastosować ten sam model:  $y = w_1x + w_b$   
 Tym razem dostosujemy go osobno dla każdego przewidywanego punktu wykorzystując informację o sąsiedztwie

Dla każdego przewidywanego  $x_i$  przypisz wagę  $\alpha_i$  dla każdego punktu danych  $x_i$  gdzie  $\tau$  służy do określenia wielkości sąsiedztwa

$$\alpha_i = e^{-\frac{(x_i - x_i)^2}{\tau}}$$

Zastanówmy się nad spodziewanym wyjściem ( $y$ ) dla tej wartości wejścia ( $x$ )

$$arg \min_w \sum_i \alpha_i \cdot (y_i - (w_1x_i + w_b))^2$$

12

---

---

---

---

---

---

---

---

### Estymator lokalnie liniowy

Dla każdego przewidywanego  $x_i$  przypisz wagę  $\alpha_i$  dla każdego punktu danych  $x_i$  gdzie  $\tau$  służy do określenia wielkości sąsiedztwa

Chcemy zastosować ten sam model:  $y = w_1x + w_b$  osobno dla każdego punktu

Metoda działa bardzo skutecznie – ale wymaga przechowania w pamięci wszystkich punktów ze zbioru danych – i wielokrotnego dopasowywania modelu – co zajmuje czas i zasoby obliczeniowe...

$$\alpha_i = e^{-\frac{(x_i - x_i)^2}{\tau}}$$

Zastanówmy się nad spodziewanym wyjściem ( $y$ ) dla tej wartości wejścia ( $x$ )

$$\arg \min_w \sum \alpha_i \cdot (y_i - (w_1x_i + w_b))^2$$

13

---

---

---

---

---

---

---

---

Model nieliniowy jest lepszy...

Skąd wiemy, że nieliniowego modelu?

Jeśli klasyfikator liniowy ma duży błąd

Musimy porównać błąd modelu liniowego i nieliniowego

Ale tutaj nie...

14

---

---

---

---

---

---

---

---

### Model wielomianowy

$$f(w, x) = b + w_1x^1 + w_2x^2 + w_3x^3 + \dots$$

Model liniowy

Model kwadratowy

To są wektory, nie pojedyncze wartości!

Modele wielomianowe działają dokładnie tak samo jak liniowe. Jedyną różnicą to konieczność ustawienia dodatkowego **metaparametru** – stopnia użytego wielomianu

Robimy to zazwyczaj poprzez zwiększanie tego stopnia aż model przestanie się poprawiać...

15

---

---

---

---

---

---

---

---

## Czy to wszystko?

*Gaussian mixture models*  
*Algorytm komitetowy* *Regresja Bayesowska*  
**Maszyna Wektorów Podtrzymujących (SVM)**  
*Multivariate Adaptive Regression Splines* *Regresja kNN*  
*Sieci konwolucyjne* *Bayesowskie sieci neuronowe*  
**Sztuczne Sieci Neuronowe (ANNs)** *Transformery*  
**Wielowarstwowe perceptrony** *Sieci Echo State*  
*Sieci radialne* **Long Short-Term Memory networks (LSTMs)**  
*Extreme learning machines*  
*Sieci rekurencyjne*

...część z nich poznamy na pozostałej części przedmiotu

16

---

---

---

---

---

---

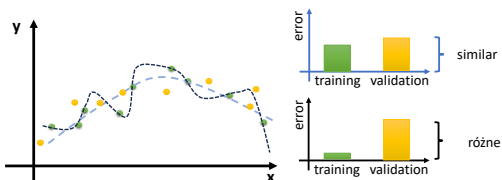
---

---

## Przeuczenie

**Przeuczenie** oznacza, że model zapamiętuje dane treningowe kosztem zdolności generalizacyjnych (zdolności uogólniania)

Rozpoznajemy ją poprzez porównanie skuteczności treningowej i skuteczności na niezależnym **walidacyjnym** podzbiore danych. Jeśli znacząco spada – model jest przeuczony.



17

---

---

---

---

---

---

---

---

## Regresja:

Chcemy zbudować model, który przewiduje **Target** na podstawie **cech**.

$$f(x)$$

*Wektor cech (wejść)*

## Klasyfikacja:

Chcemy zbudować model, który przewiduje **Etykię** na podstawie **cech**.

18

---

---

---

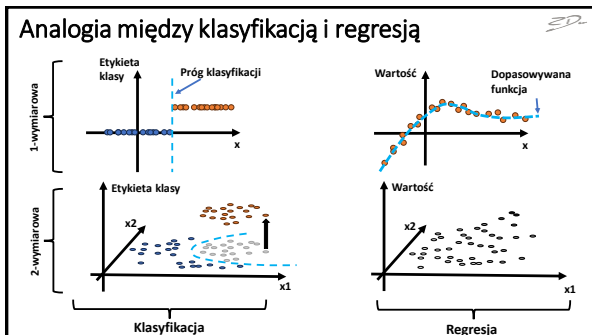
---

---

---

---

---



19

---

---

---

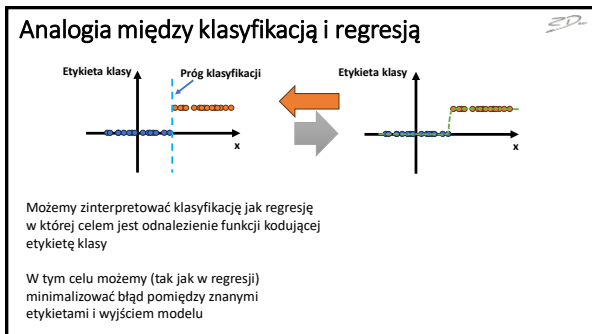
---

---

---

---

---



20

---

---

---

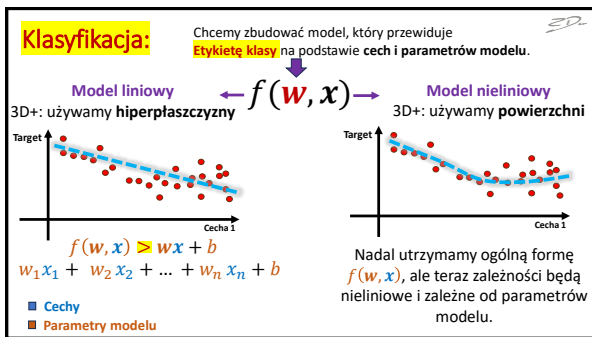
---

---

---

---

---



21

---

---

---

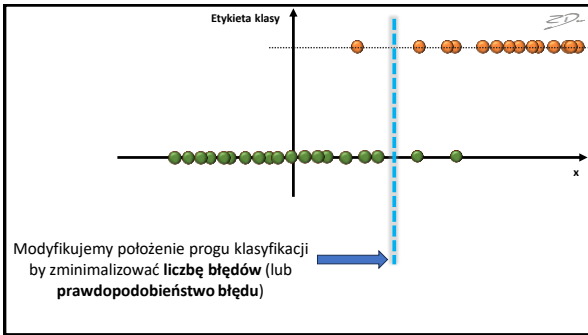
---

---

---

---

---



22

---

---

---

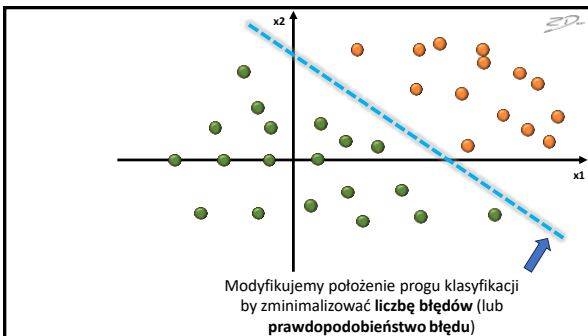
---

---

---

---

---



23

---

---

---

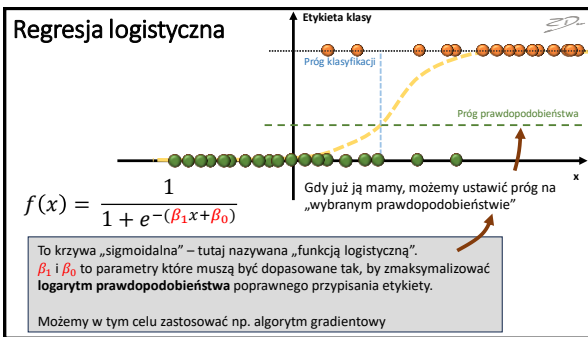
---

---

---

---

---



24

---

---

---

---

---

---

---

---



## Regresja logistyczna

- LR to klasyfikacyjny odpowiednik regresji liniowej
- Dobrze działa również w przestrzeni wielowymiarowej (mamy jedynie więcej parametrów do nauczenia)
- Może wpasować „klasyfikator liniowy” w sposób „najlepszy z możliwych” (tzn. klasyfikacja ma interpretację probabilistyczną)
- Nie jest w stanie rozwiązywać problemów nieseparalnych liniowo



25

---

---

---

---

---

---

---

---

## Materiał do powtórki:

1. Wyjaśnij na przykładzie zadanie regresji (w tym wielowymiarowej)
2. Wyjaśnij algorytmy regresji liniowej, wielomianowej, estymatora lokalnie liniowego. Podaj ich właściwości (wady, zalety)
3. Wyjaśnij czym jest przeuczenie i jak można je wykryć
4. Wyjaśnij czym się różni klasyfikacja od regresji (w tym na przykładach)
5. Wyjaśnij zasadę działania regresji logistycznej, podaj jej zalety i wady i podaj wzór na funkcję logistyczną

26

---

---

---

---

---

---

---

---