



Politechnika
Wroclawska

Algorytmy sztucznej inteligencji w Przemysle 4.0

Uczenie nadzorowane

Dr inż. Radosław Idzikowski



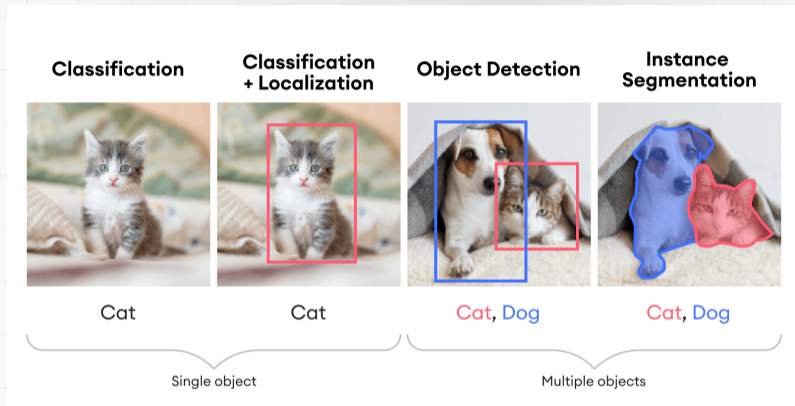
HR EXCELLENCE IN RESEARCH



Zadania w Uczeniu Nadzorowanym

- ▶ **Regresja**
 - ▶ Liniowa
 - ▶ Wielomianowa
- ▶ **Drzewa decyzyjne**
- ▶ **Lasy losowe**
- ▶ **Klasyfikacja**
 - ▶ KNN
 - ▶ Trees
 - ▶ Regresja logiczna
 - ▶ Nive-Bayes
 - ▶ SVM

Zadania w Uczeniu Nadzorowanym



Rysunek 1: Różnice pomiędzy wybranymi problemami widzenia komputerowego (ang. *computer vision*)



Generalizacja

Trening sieci neuronowych wiąże się z wyzwaniem osiągnięcia efektywnej generalizacji — zdolności modelu do skutecznego radzenia sobie z nowymi, nieznanymi danymi. To kluczowe, biorąc pod uwagę ograniczoną ilość danych treningowych i potencjalne błędy w nich. Głównym celem jest stworzenie modelu, który nie tylko działa skutecznie na zbiorze treningowym, ale także na nowych danych.



Zbiór danych

treningowy

Używany w procesie treningu do modyfikacji wag sieci neuronowej, uczenie nadzorowane ma na celu, aby sieć dopasowywała dane wejściowe do oczekiwanych rezultatów. Proces ten polega na indukowaniu ogólnych zasad z konkretnych przykładów. Aby osiągnąć zdolność do generalizacji, zbiór treningowy powinien być zróżnicowany i zawierać reprezentatywne przykłady. Jedna "epoka" treningu oznacza jedno pełne przejście przez cały zbiór treningowy. Rozmiar tego zbioru często wynosi od 60% do 80% całej dostępnej próbki danych.

walidacyjny

równocześnie używany z zbiorem treningowym, służy do oceny jakości modelu podczas treningu. Jest kluczowy do dostrojenia hiperparametrów i konfiguracji modelu oraz do monitorowania poprawności procesu treningu, oceniając kierunek jego postępu i wykrywając ewentualne błędy. Używany po każdej epoce, pozwala na ocenę jakości modelu na danych niewykorzystanych wprost do treningu. Typowy rozmiar zbioru walidacyjnego to od 10% do 20% całej dostępnej próbki danych.

testowy

Zbiór testowy, używany do końcowej oceny modelu po treningu, dostarcza informacji o jego jakości na danych nie użytych wcześniej. Stanowi symulację przyszłego zachowania modelu w warunkach produkcyjnych. Rozmiar zbioru testowego często wynosi od 10% do 20% całej dostępnej próbki danych.



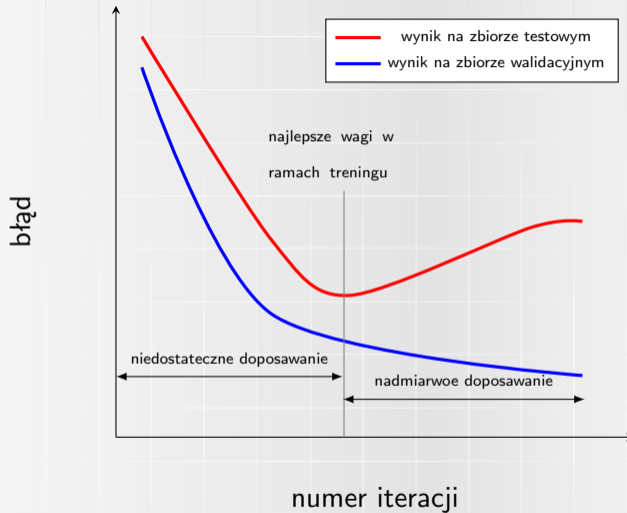
Generalizacja w Uczeniu Maszynowym

- ▶ Uczenie maszynowe wyróżnia się przez uwzględnienie generalizacji, różniąc się tym od klasycznej optymalizacji.
- ▶ Minimalizacja błędu na zbiorze treningowym nie jest celem ostatecznym. Priorytetem jest wynik na zbiorze testowym, który służy jako podstawa do oceny przydatności modelu na nowych danych.

Warunki

- ▶ Poprawa wyników na zbiorze testowym wymaga pewnych założeń, takich jak niezależność danych treningowych, walidacyjnych i testowych, oraz wspólny rozkład prawdopodobieństwa.
- ▶ Założenie o wspólnym rozkładzie umożliwia powiązanie błędu na zbiorze treningowym z błędem na zbiorze testowym.

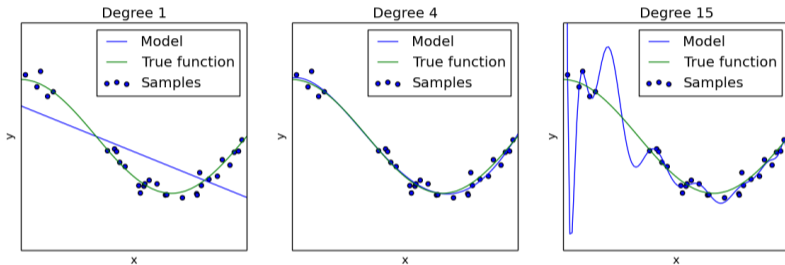
Underfitting a Overfitting



Rysunek 2: Przebieg procesu trenowania sieci neuronowej

Underfitting a Overfitting

Regresja liniowa



Rysunek 3: Ten przykład ilustruje problemy z niedostosowaniem i nadmiernym dostosowaniem oraz sposób korzystania z regresji liniowej z cechami wielomianowymi do przybliżania funkcji nieliniowych



Niedostateczne Dopasowanie Modelu

- ▶ Niedostateczne dopasowanie występuje, gdy model osiąga słabe wyniki już na zbiorze treningowym, co uniemożliwia generalizację.
- ▶ Przyczyny:
 - ▶ Niewielki rozmiar danych treningowych.
 - ▶ Brak odpowiedniego przygotowania lub słaba jakość danych.
 - ▶ Niedostateczna złożoność modelu dla rozwiązywanego problemu.
- ▶ Możliwe rozwiązania:
 - ▶ Inżynieria cech poprzez dostosowanie danych wejściowych.
 - ▶ "Czyszczenie" danych eliminując błędy i redukując szumy.
 - ▶ Wydłużenie procesu treningu.
 - ▶ Użycie bardziej skomplikowanego modelu.



Nadmierne Dopasowanie Modelu

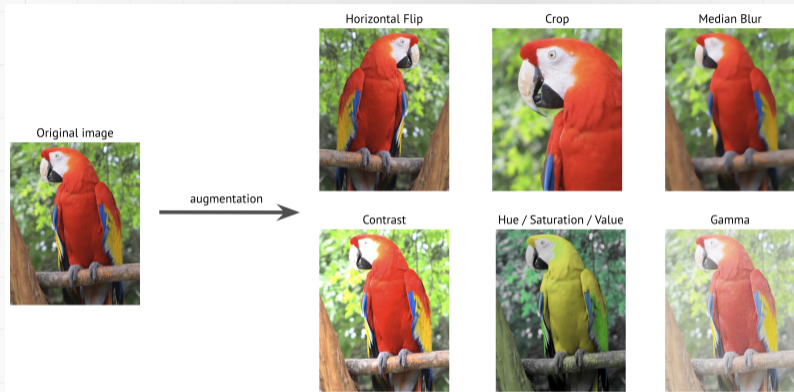
- ▶ Nadmierne dopasowanie występuje, gdy model uzyskuje dobre wyniki na zbiorze treningowym, ale pogarsza się na zbiorze testowym, co prowadzi do zmniejszenia zdolności do generalizacji.
- ▶ Przyczyny:
 - ▶ Wykorzystanie zbyt złożonego modelu dla rozwiązania danego problemu.
 - ▶ Niewystarczający rozmiar danych treningowych.
- ▶ Możliwe rozwiązania:
 - ▶ Zwiększenie rozmiaru danych treningowych.
 - ▶ Zmniejszenie złożoności modelu.
 - ▶ Zastosowanie technik zapobiegających nadmiernemu dopasowaniu.



Techniki Przeciwdziałania Nadmiernemu Dopasowaniu

- ▶ **Wcześniejsze Przerwanie Treningu (Early Stopping):** Używanie wag modelu z poprzednich iteracji, zanim wyniki na zbiorze walidacyjnym ulegną pogorszeniu.
- ▶ **Walidacja Krzyżowa:** Podział danych na części do treningu i walidacji, eliminując problem jednorodności zbioru treningowego.
- ▶ **Metody Regularyzacji, np. Dropout:** Losowe wyłączenie neuronów podczas treningu, zmniejszając złożoność modelu.
- ▶ **Augmentacja:** Zmiany w danych treningowych, np. w przetwarzaniu obrazu, aby uzyskać nowe próbki, zapobiegając nadmiernemu dopasowaniu.

Agumentacja



Rysunek 4: Augmentacja zdjęcia jako przykład sztucznego rozszerzania zbioru danych treningowych

- ▶ *Data Leakage* to sytuacja, gdzie dane treningowe zawierają informacje ułatwiające modelowi osiągnięcie bardzo dobrych wyników na danej próbce, ale z obserwowalnym spadkiem jakości na nowych danych.
- ▶ Przykłady wycieku danych: Dane czasowe, np. przewidywanie kursu giełdowego na podstawie danych z ostatniego miesiąca, gdzie część danych do przewidywania jest dostępna w zbiorze treningowym.



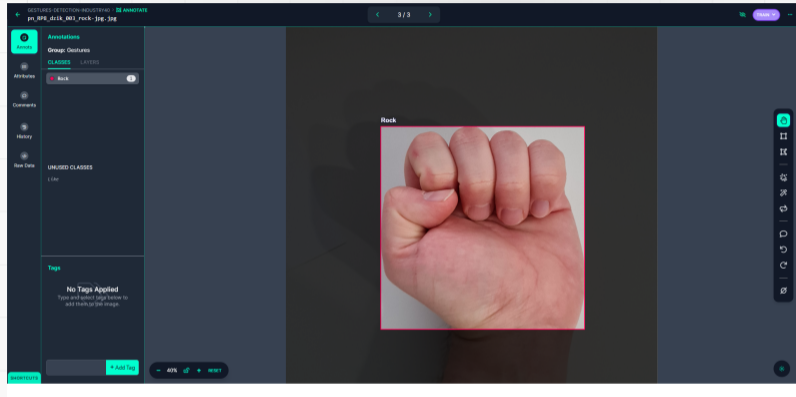
Proces Uczenia

Trening modelu

1. Zebranie danych.
2. Przygotowanie danych treningowych.
3. Dopasowanie modelu do danych treningowych.

Bounding Box

prostokątna ramka określająca granice obszaru zawierającego obiekt na obrazie. W kontekście *computer vision* i przetwarzania obrazu, *bounding box* jest używany do wskazania lokalizacji i zakresu obiektu, co umożliwia modelom uczenia maszynowego identyfikację i śledzenie obiektów na obrazie. Składa się z czterech współrzędnych definiujących lewą górną oraz prawy dolny narożnik,








Rysunek 5: Roboflow

YOLO v5

(You Only Look Once) to efektywny model detekcji obiektów w czasie rzeczywistym.

- ▶ Jednoetapowy proces detekcji obiektów.
- ▶ Wydajność na poziomie state-of-the-art.
- ▶ Zastosowanie w różnych dziedzinach, w tym w rozpoznawaniu obrazów medycznych czy przetwarzaniu wideo.

Nano	Small	Medium	Large	XLarge
				
YOLOv5n	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x
4 MB _{FP16} 6.3 ms _{V100} 28.4 mAP _{COCO}	14 MB _{FP16} 6.4 ms _{V100} 37.2 mAP _{COCO}	41 MB _{FP16} 8.2 ms _{V100} 45.2 mAP _{COCO}	89 MB _{FP16} 10.1 ms _{V100} 48.8 mAP _{COCO}	166 MB _{FP16} 12.1 ms _{V100} 50.7 mAP _{COCO}

Wersje YOLO

- ▶ YOLOv2 (2016) wprowadził ulepszenia, takie jak batch normalization, anchor boxes i dimension clusters.
- ▶ YOLOv3 (2018) dodał bardziej efektywną sieć podstawową, multiple anchors i spatial pyramid pooling.
- ▶ YOLOv4 (2020) wprowadził innowacje, takie jak Mosaic data augmentation, anchor-free detection head i nowa funkcja straty.
- ▶ YOLOv5 dalej poprawił wydajność modelu i dodano nowe funkcje, np. optymalizację hiperparametrów i automatyczny eksport do popularnych formatów.
- ▶ YOLOv6 został udostępniony przez Meituan w 2022 roku i jest używany w wielu autonomicznych robotach dostawczych firmy.
- ▶ YOLOv7 dodaje zadania takie jak szacowanie pozy na zbiorze danych COCO keypoints.
- ▶ YOLOv8 to najnowsza wersja YOLO stworzona przez Ultralytics. Jako nowoczesny model state-of-the-art (SOTA), wprowadza nowe funkcje i ulepszenia dla lepszej wydajności, elastyczności i efektywności.



Zalety YOLOv5

- ▶ **Wysoka Wydajność:** YOLOv5 oferuje szybką detekcję obiektów, co sprawia, że jest efektywny nawet w czasie rzeczywistym.
- ▶ **Skuteczność:** Model ten charakteryzuje się wysoką dokładnością detekcji, co przyczyniło się do jego popularności.
- ▶ **Łatwa Implementacja:** Dostępność jako otwarte oprogramowanie ułatwia integrację i dostosowanie YOLOv5 do różnych zastosowań.
- ▶ **Nowoczesne Funkcje:** YOLOv5 wprowadza innowacyjne funkcje, takie jak optymalizacja hiperparametrów i automatyczny eksport do popularnych formatów.
- ▶ **Wszechstronność:** Model ten obsługuje pełen zakres zadań wizji sztucznej, w tym detekcję, segmentację, szacowanie pozy, śledzenie i klasyfikację.

Miary oceny klasyfikatora

Macierz pomyłek (*Confusion matrix*)

Rozważmy klasyfikator binarny

	"A"	"B"
Predicted as "A"	True positive (TP)	False positive (FP)
Predicted as "B"	False negative (FN)	True negative (TN)

$$P = TP + FN \quad (1)$$

$$N = FP + TN \quad (2)$$



Miary oceny klasyfikatora

Trafność

$$\text{accuracy} = \frac{TP + TN}{P + N} \quad (3)$$

	"A"	"B"
Predicted as "A"	25	18
Predicted as "B"	7	50

$$\text{accuracy} = \frac{25 + 50}{100} = 0,75 \quad (4)$$

	"A"	"B"
Predicted as "A"	0	20
Predicted as "B"	0	980

$$\text{accuracy} = \frac{0 + 980}{1000} = 0,98 \quad (5)$$



Miary oceny klasyfikatora

Pokrycie

$$\text{recall} = TPR = \frac{TP}{TP + FN} \quad (6)$$

	"A"	"B"
Predicted as "A"	25	18
Predicted as "B"	7	50

$$\text{recall} = \frac{25}{25 + 18} = 0,58 \quad (7)$$

	"A"	"B"
Predicted as "A"	0	20
Predicted as "B"	0	980

$$\text{recall} = \frac{0}{0 + 20} = 0 \quad (8)$$



Miary oceny klasyfikatora

Dokładność

$$\text{precision} = \frac{TP}{TP + FP} \quad (9)$$

	"A"	"B"
Predicted as "A"	25	18
Predicted as "B"	7	50

$$\text{precision} = \frac{25}{25 + 7} = 0,781 \quad (10)$$

	"A"	"B"
Predicted as "A"	50	20
Predicted as "B"	50	880

$$\text{precision} = \frac{50}{50 + 50} = 0,5 \quad (11)$$



Miary oceny klasyfikatora

Dokładność vs Pokrycie

Obie miary w skrajnych przypadkach mogą zwracać rozbieżne wyniki.

	"A"	"B"
Predicted as "A"	50	0
Predicted as "B"	950	0

$$\text{precision} = 0,05 \quad (12)$$

$$\text{recall} = 1 \quad (13)$$

	"A"	"B"
Predicted as "A"	3	47
Predicted as "B"	0	950

$$\text{precision} = 1 \quad (14)$$

$$\text{recall} = 0,06 \quad (15)$$



Miary oceny klasyfikatora

F1

$$F_1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (16)$$

	"A"	"B"
as "A"	50	0
as "B"	950	0

$$\text{precision} = 0,05 \quad (17)$$

$$\text{recall} = 1 \quad (18)$$

$$F_1 = 0,05 \quad (19)$$

	"A"	"B"
as "A"	3	47
as "B"	0	950

$$\text{precision} = 1 \quad (20)$$

$$\text{recall} = 0,06 \quad (21)$$

$$F_1 = 0,6 \quad (22)$$

	"A"	"B"
as "A"	34	16
as "B"	7	43

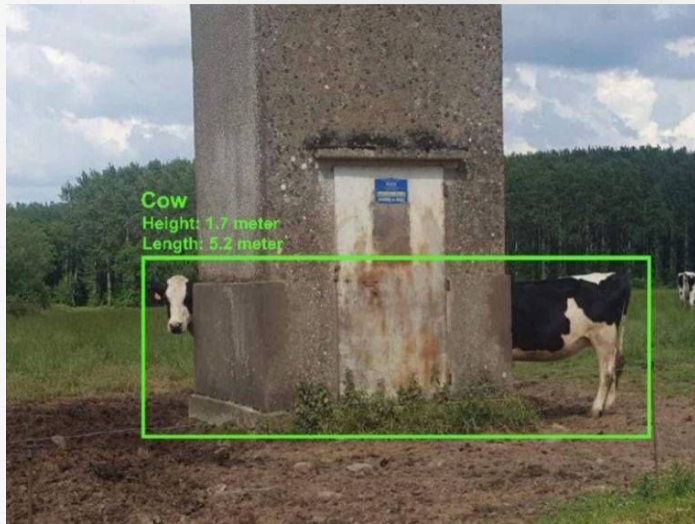
$$\text{precision} = 0,83 \quad (23)$$

$$\text{recall} = 0,68 \quad (24)$$

$$F_1 = 0,74 \quad (25)$$



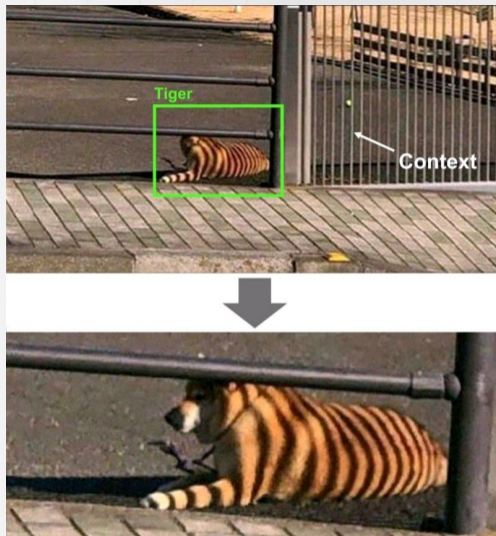
Czy kontekst jest istotny?



Rysunek 6: Błędnie sklasyfikowana krowa



Czy kontekst jest istotny?



Rysunek 7: Błędnie sklasyfikowany tygrys



Zastosowania w Przemysle 4.0

- ▶ **Monitorowanie Maszyn:** pozwala na wczesne rozpoznawanie wzorców awarii i optymalizację utrzymania maszyn.
- ▶ **Prognozowanie Zużycia Narzędzi:** może przewidywać zużycie narzędzi w obrabiarkach, umożliwiając planowanie konserwacji.
- ▶ **Monitorowanie Jakości Produkcji:** używane do monitorowania i poprawy jakości produkcji.
- ▶ **Optymalizacja Procesów Logistycznych:** Dane związane z dostawami i magazynowaniem są analizowane w celu optymalizacji procesów logistycznych.