

Inżynieria Mechatroniczna

Podstawy Sztucznej Inteligencji i Ucznienia Głębokiego:
5: Dane i modele – od strony praktycznej

Ziemowit Dworakowski
AGH w Krakowie

1

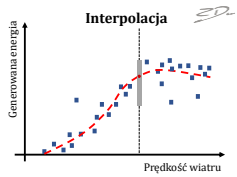
SD

Ogólna pętla przetwarzania danych i testowania modeli

2

Różnica pomiędzy klasycznym przetwarzaniem danych i tym opartym na AI

Klasycznie:



3

Źródła danych

Timestamp („Znacznik czasu”)
 Temperatura zewnętrzna
 Amplituda drgań

f_1
 f_2
 f_3
 f_4
 f_n

Zbieramy:

- Wektory zmiennych (cech)
- Sygnały 1D (np. czasowe)
- Sygnały 2D (np. obrazy)
- Metadane i zmienne opisowe

7

Źródła danych

Nie ma znaczenia, czy pracujemy z sygnałami, danymi SCADA, widmami, obrazami, sygnaturami drgań... Ostatecznie i tak musimy je zamienić na wektory cech.

Zbieramy:

- Wektory zmiennych (cech)
- Sygnały 1D (np. czasowe)
- Sygnały 2D (np. obrazy)
- Metadane i zmienne opisowe

Ekstrakcja cech

Cechy to jedyny typ danych który może być używany przez płytkie modele AI

8

Ekstrakcja cech

Które cechy zawierają znaczącą informację?

f_1
 f_2
 f_3
 f_4
 f_n

Idealna cecha powinna być:

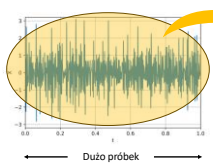
- Wrażliwa wyłącznie na wartości docelowe
- Niewrażliwa na inne zmienne i warunki eksperymentalne

Zwykle trudno takie znaleźć

9

Ekstrakcja cech

SD



- f1
- f2
- f3
- f4
- ...
- fn

Które cechy zawierają znaczącą informację?

- Akceptowalna cecha powinna:**
- Być wrażliwa na niektóre wartości docelowe
 - Być wrażliwa na warunki eksperymentalne – pozwalając na ich filtrację
 - Nie być zbyt zasumiona
 - Zawierać jakąś unikatową informację

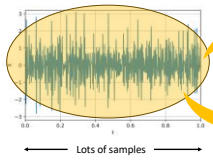
To nie wygląda jak jasny przepis na poszukiwanie cech...

10

Ekstrakcja cech

SD

Zamiast tego użyjmy jednego z poniższych sposobów:



Wyznamy więcej cech niż potrzebujemy a potem dokonamy **selekcji** podzbioru który wydaje się działać najlepiej

To wygląda jak problem optymalizacyjny...

Pomyślimy, jak można rozpoznać klasy „ręcznie” a potem zastosujemy cechy, które na to pozwalają

a to wymaga wiedzy dziedzinowej...

Tak czy owak – będziemy wykonywali dużo zaawansowanego przetwarzania sygnałów, by uzyskać cechy które faktycznie pozwalają na klasyfikację

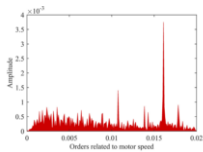
11

Ekstrakcja cech

SD

Wykonajmy kilka przykładów:

Zapis drgań w czasie – do rozpoznania stanu technicznego maszyny



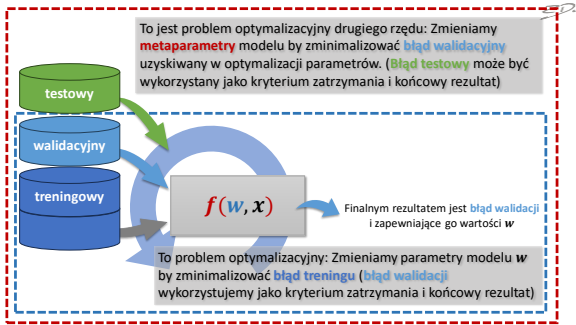
RMS zawiera informację o energii zawartej w sygnale

Kurtoza informuje jak bardzo „poszarpany” jest sygnał

Wyznaczenie widma częstotliwościowego albo rzędów pozwala na wyznaczenie energii w pasmach, która może korelować z konkretnymi typami uszkodzeń

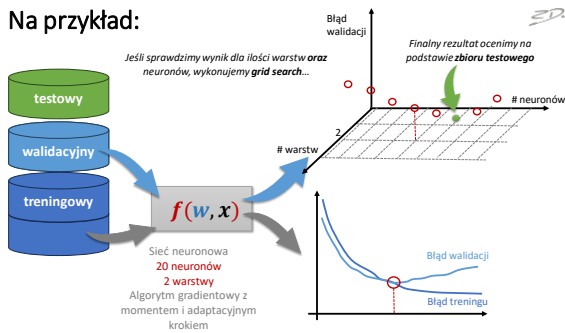
Jeśli poszukujemy konkretnych uszkodzeń, możemy zaprojektować cechy dostosowane do ich wykrywania.

12



16

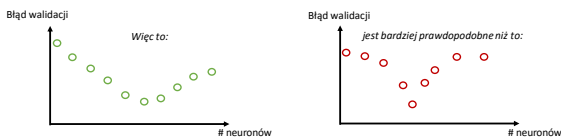
Na przykład:



17

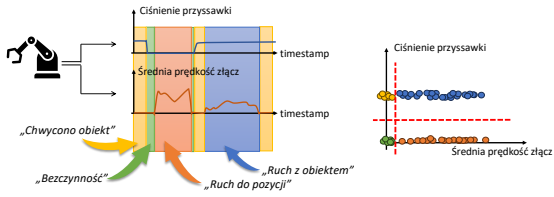
Właściwości optymalizacji metaparametrów

- Zazwyczaj tylko niektóre metaparametry są naprawdę istotne (Możemy użyć relatywnie prostych algorytmów optymalizacyjnych – np. *grid search*)
- Zwykle wartości domyślne są wystarczająco dobre... (w większości praktycznych zastosowań można przy nich pozostać)
- Funkcje celu zwykle są stosunkowo „łagodne” – nie zawierają „stromych” minimum (Eksploracja minimum nie jest potrzebna, wystarczy znaleźć się w pobliżu minimum)



18

Zastosowanie w rozpoznawaniu stanu pracy *SD*

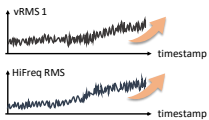


Możemy oczywiście dysponować większą ilością zmiennych operacyjnych (tworząc wielowymiarową przestrzeń cech). Tak czy owak, **To zadanie klasyfikacyjne – które umiemy rozwiązać!**

25

Zastosowanie w predykcji (np. RUL) *SD*

Powiedzmy, że mamy maszynę wirnikową, dla której mierzymy takie cechy:



Wzrost wartości vRMS oraz drgań w wysokich częstotliwościach sugeruje zużywanie się łożysk...

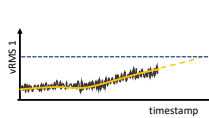
Więc pewnie ulegną uszkodzeniu...

Kiedyś...

Ale kiedy?

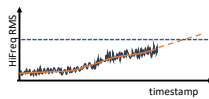
26

Zastosowanie w predykcji (np. RUL) *SD*



Możemy sprawdzić na jakim poziomie wartości cech zwykle następuje uszkodzenie a następnie wpisać w punkty jakiś model przewidujący trend – i sprawdzić kiedy linie się przetną...

To zadanie regresji, umiemy takie rozwiązywać!



Lub nawet lepiej:

Możemy zebrać dane łączące wartości mierzonych cech z pozostałym czasem życia (Remaining Useful Life, RUL) łożyska – i potem wykonać zadanie regresji przewidując ten czas

Co również umiemy zrobić!

27

Zastosowanie w modelowaniu procesów

SD



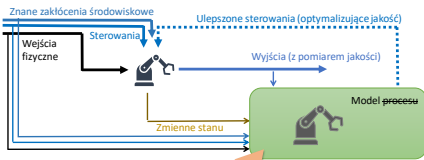
Dobry model powinien być w stanie odpowiedzieć na pytanie „Jeśli wejścia mają wartość x , jakie będzie wyjście...?”

A to też zadanie regresji, które umiemy rozwiązywać!

28

Zastosowanie w identyfikacji

SD



Dobry model powinien być w stanie odpowiedzieć na pytanie „Jeśli wejścia mają wartość x , jakie będzie wyjście...?”

A to też zadanie regresji, które umiemy rozwiązywać!

29

Jednoklasowa klasyfikacja (detekcja anomalii)

SD

W niektórych problemach mamy nadmiar obiektów należących do jednej klasy i znaczny niedobór tych należących do innej:

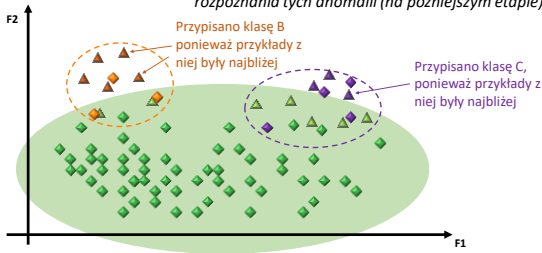
- „**Zdrowi** ludzie”
- „**Normalne** zachowanie w miejscu publicznym”
- „**Normalny** stan operacyjny linii montażowej
- „**Typowa** pogoda we wrześniu”

Wówczas naszym celem może być modelowanie zakresu **normalnego** i wykrywanie wszystkich odstępstw od normy (**anomalii, outlierów**)

W tym zadaniu zwykle „nie wiemy czego oczekiwać” – tzn. istnieje potencjalnie nieskończony zbiór „odstępstw od normy”

30

1. Naucz się „zakresu normalnego”
2. Wykryj, co wykracza poza „zakres normalny”
3. Jeśli masz przykłady innych klas, możesz ich użyć do rozpoznania tych anomalii (na późniejszym etapie)



31

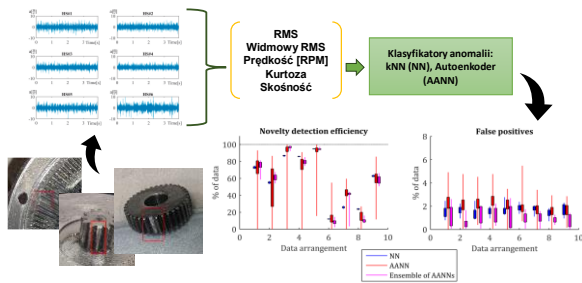
Jednoklasowa klasyfikacja (detekcja anomalii)

Jak zdefiniować „zakres normalny”?

1. Możemy określić maksymalny dystans pomiędzy próbkami w zbiorze normalnym i wykrywać punkty, które go znacząco przekraczają > np. **kNN**
2. Możemy estymować gęstość prawdopodobieństwa obecności próbek w przestrzeni cech – a potem wykrywać próbki, które znajdują się w obszarach zerowego prawdopodobieństwa -> np. **GMM**
3. Możemy wytrenować system regresyjny by przewidywał jedną cechę na podstawie pozostałych. Dla znanych danych (w zakresie normalnym) powinien działać skutecznie, dla nietypowych danych (anomalii) będzie generował duże błędy -> np. **sieć MLP** lub **autoenkoder**

32

Zastosowanie w wykrywaniu anomalii



33

(Ponieważ działają na podstawie zbioru danych pozyskanych z jakiegos rzeczywistego rozkładu. Przy ponownym pozyskaniu danych rezultaty będą inne!)

SD

Klasyfikatory i regresory powinny być **zawsze** traktowane jako niedeterministyczne

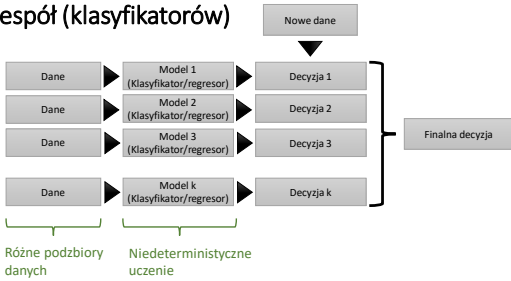


Jak zachować wysoką powtarzalność?
(Jak się upewnić, że wyniki które uzyskujemy nie wynikają z przypadkowego szczęścia lub pecha w rozkładzie danych?)

34

Zespół (klasyfikatorów)

SD



Im bardziej **różnorodne** i **dokładne** modele, tym lepiej

35

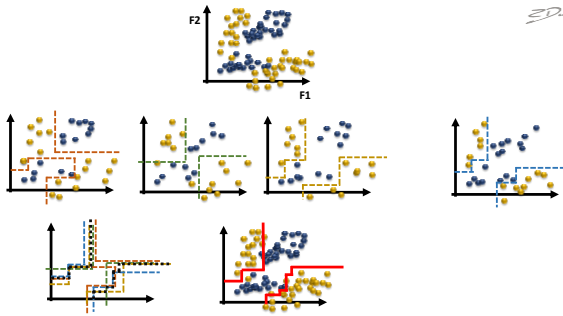
Las losowy



1. Wykonaj agregację bootstrapową (**Bagging, Bootstrap Aggregating**) poprzez losowanie (ze zwracaniem) wielu małych podzbiorów danych uczących
2. Jeśli zbiór danych zawiera dużo cech, wykonaj losowanie również w ich zakresie (tzn. niech różne podzbiory korzystają z przypadkowego podzbioru cech)
3. Wytrenuj proste drzewa decyzyjne na tych podzbiorach (każde drzewo używa innego podzbioru)
4. Uśrednij odpowiedzi z wszystkich drzew

Im bardziej **różnorodne** i **dokładne** modele, tym lepiej

36



SD

37

Zespoły klasyfikatorów

SD

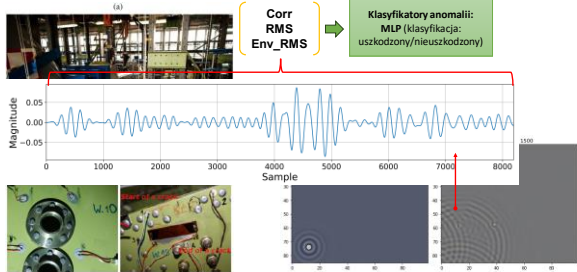
Klasyfikacja zespołowa zapewnia zwiększoną stabilność (powtarzalność) kosztem:

- braku bezpośredniej kontroli nad procesem klasyfikacji,
- braku możliwości optymalizacji struktury klasyfikatorów, oraz
- dużego obciążenia pamięci i dużej mocy obliczeniowej

38

Zespoły klasyfikatorów: przykład

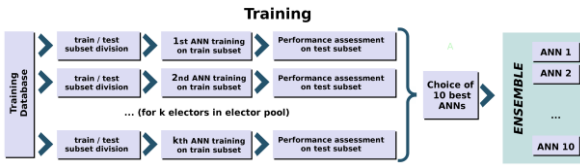
SD



39

Zespoły klasyfikatorów: przykład

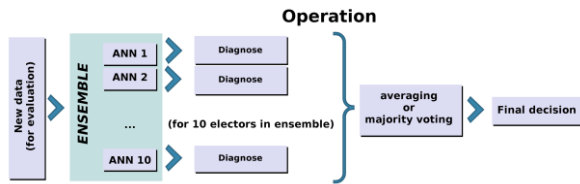
SD



40

Zespoły klasyfikatorów: przykład

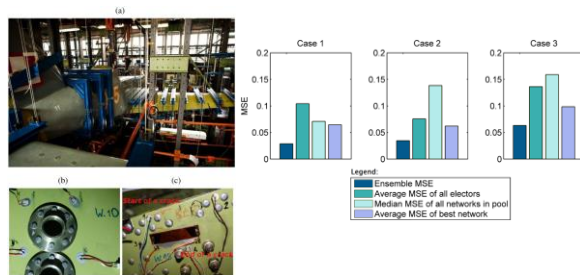
SD



41

Zespoły klasyfikatorów: przykład

SD



42

Materiał do powtórki:

SD

1. Jak działa drzewo decyzyjne i jakie ma cechy?
2. Jakie są źródła danych w inżynierii i jak można je przetwarzać w celu wydobycia cech (na przykładach)?
3. Zaprezentuj schemat treningu klasyfikatora ze wskazaniem celów dwóch poziomów optymalizacji
4. Jakie są charakterystyczne właściwości problemu optymalizacji metaparametrów i jakie są sposoby jego rozwiązywania?
5. Jakie miary jakości są wykorzystywane w klasyfikacji i regresji, jak działa i do czego służy tablica pomyłek?
6. Przedstaw przykłady praktycznego stosowania metod uczenia maszynowego w identyfikacji, predykcji, rozpoznawaniu stanów operacyjnych, modelowaniu procesów i detekcji anomalii.
7. Jak działa klasyfikacja zespołowa (Na przykładzie lasu losowego), jakie ma wady i zalety?
