

Eksploracja danych

ELEMENTY INŻYNIERII WEJŚCIA I WYJŚCIA

Wojciech Waloszek

wowal@eti.pg.gda.pl

Teresa Zawadzka

tegra@eti.pg.gda.pl

Katedra Inżynierii Oprogramowania

Wydział Elektroniki, Telekomunikacji i Informatyki

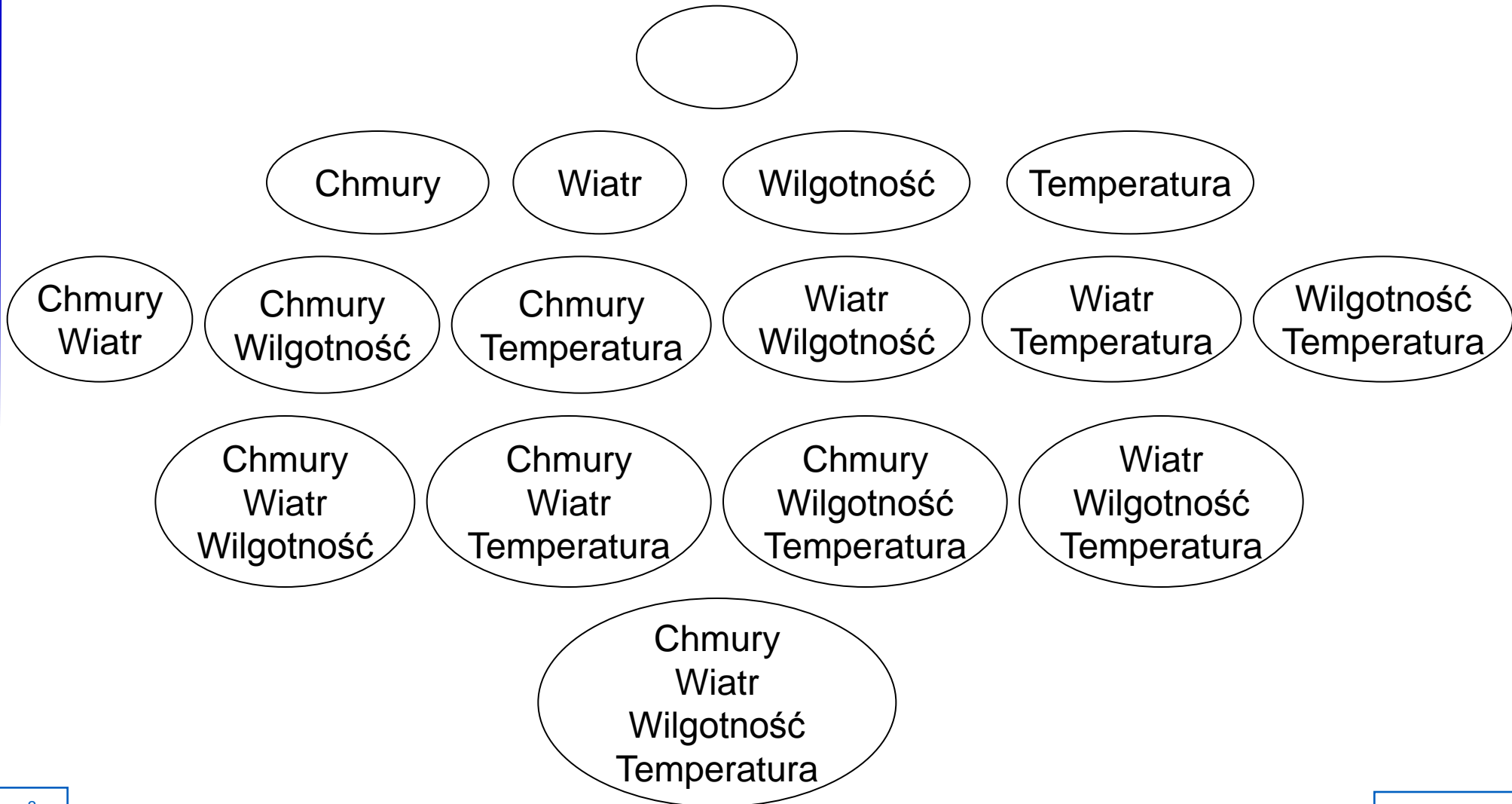
Politechnika Gdańska



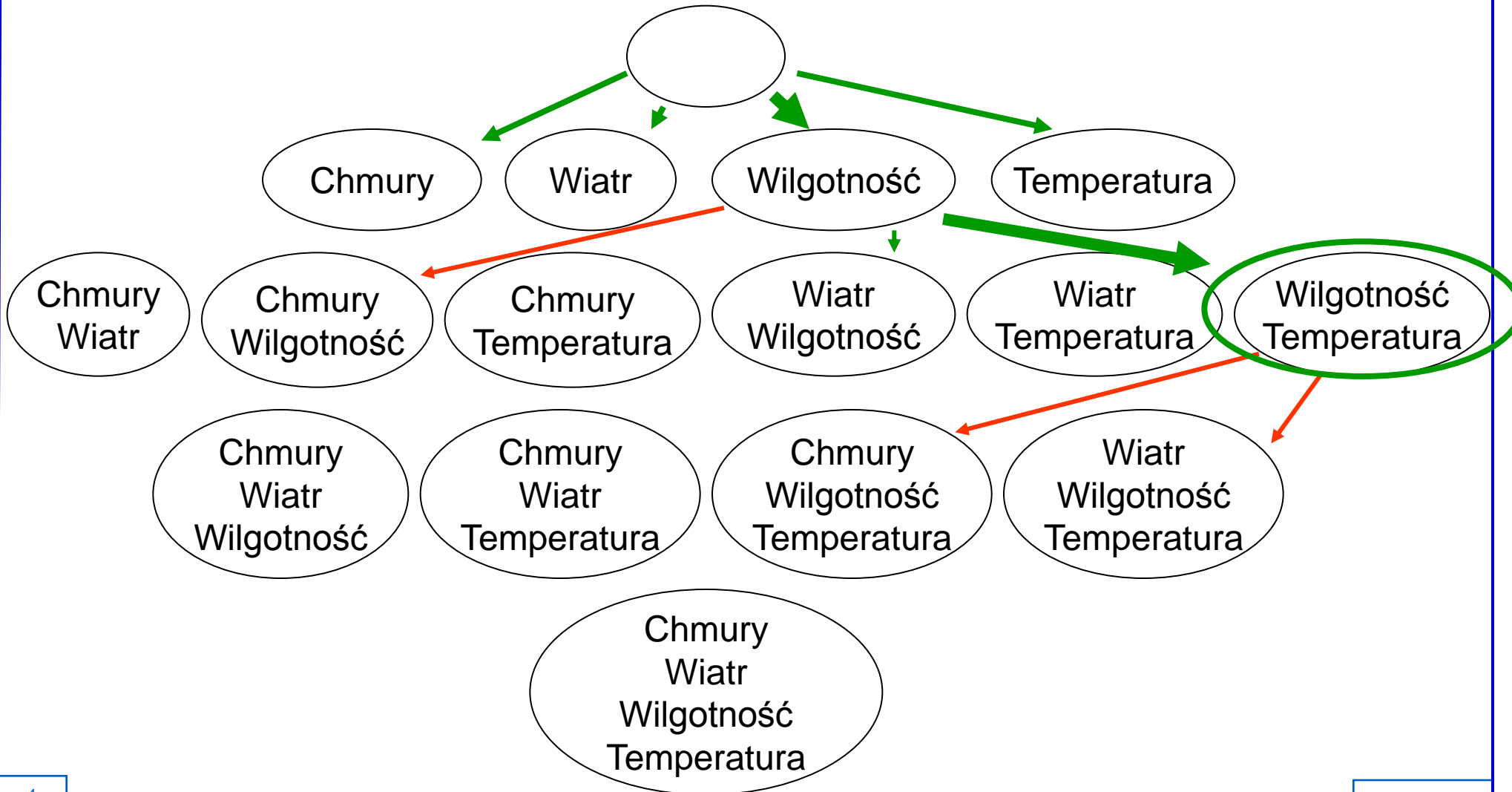
Techniki doboru atrybutów

- Ze względu na zazwyczaj zachłanny charakter metod tworzenia klasyfikatorów, zmniejszenie liczby atrybutów może często doprowadzić do uzyskania lepszych efektów,
- Stosuje się dobór atrybutów *naprzód* i *wstecz* metodą zachłanną.

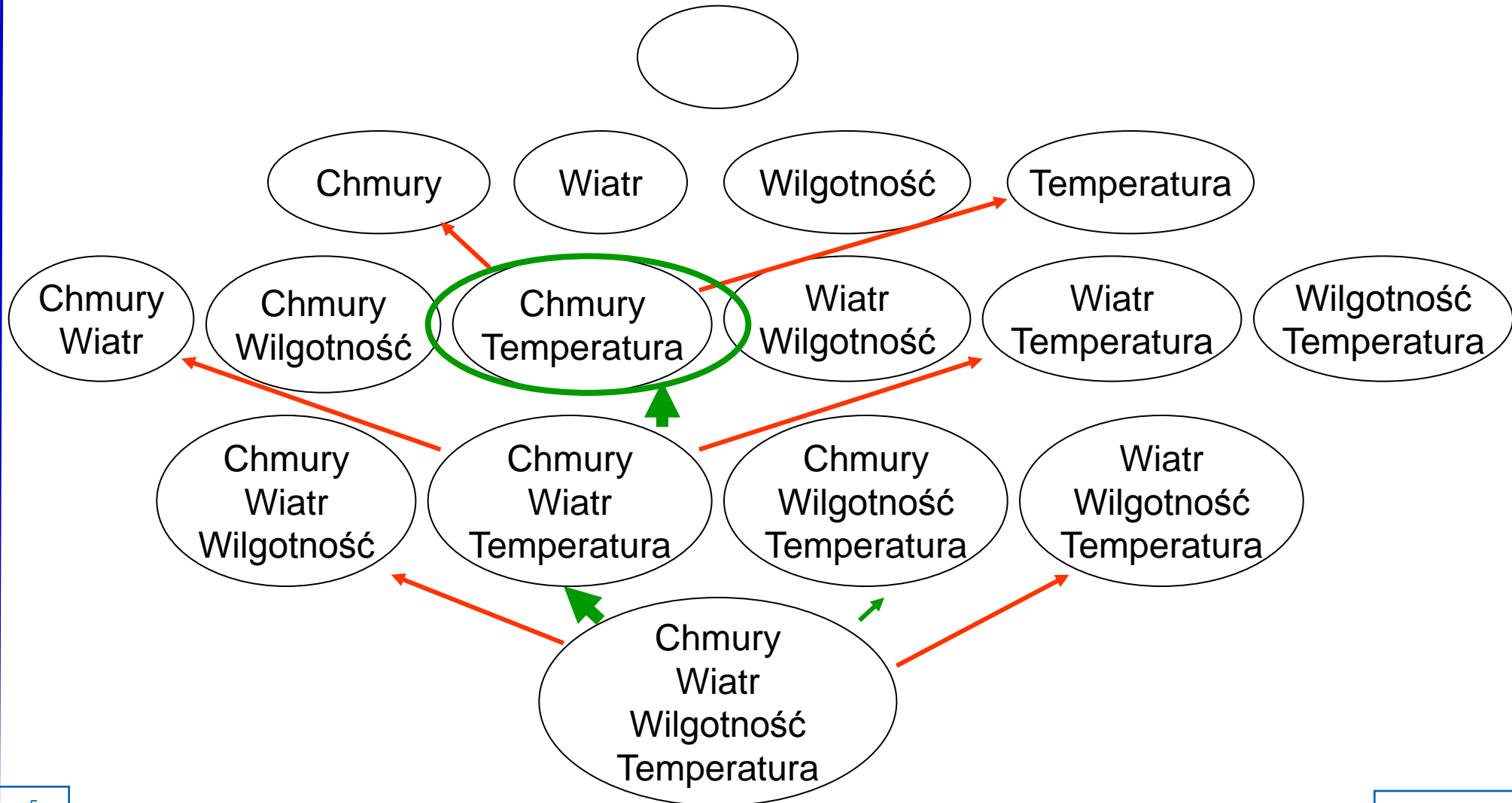
Przestrzeń atrybutów



Dobór atrybutów *naprzód*



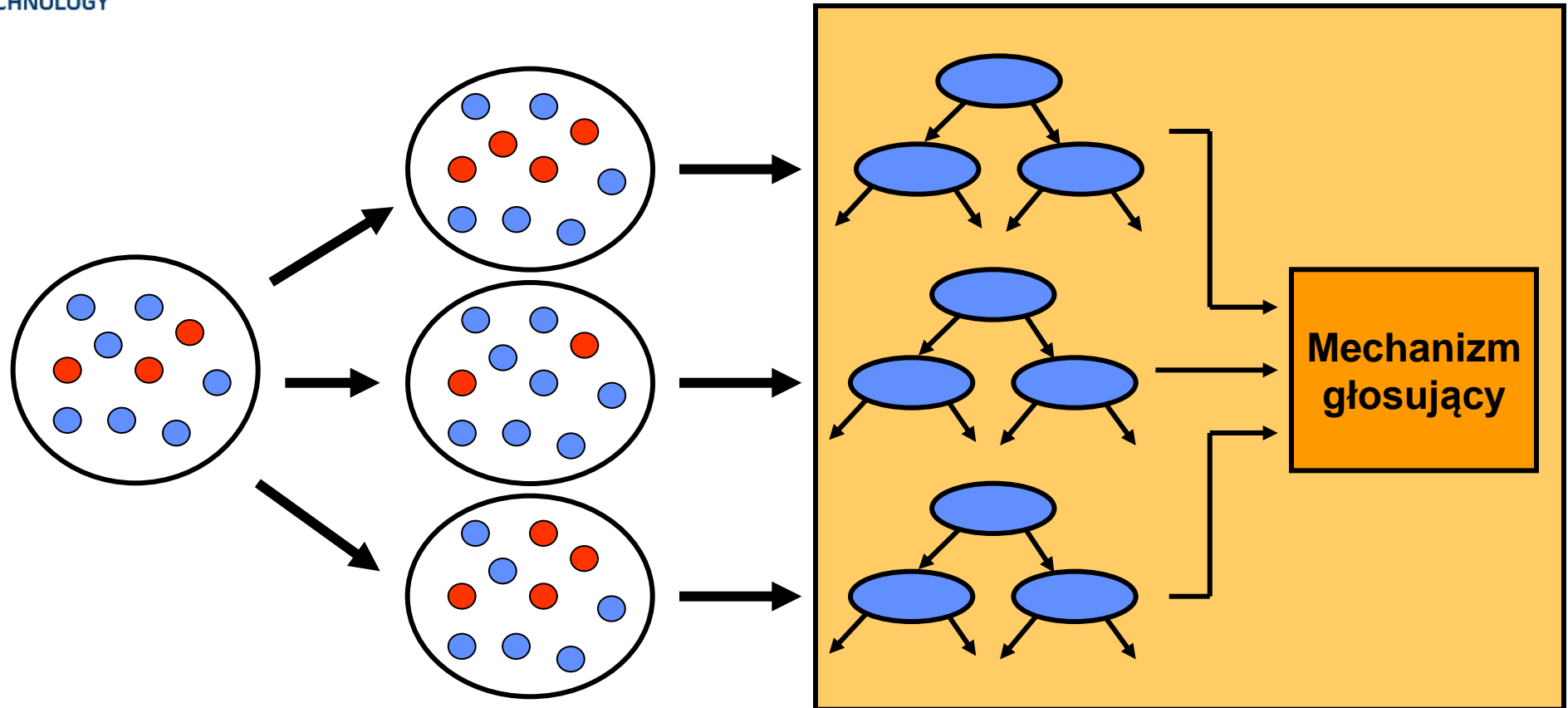
Dobór atrybutów *wstecz*



Łączenie kilku klasyfikatorów

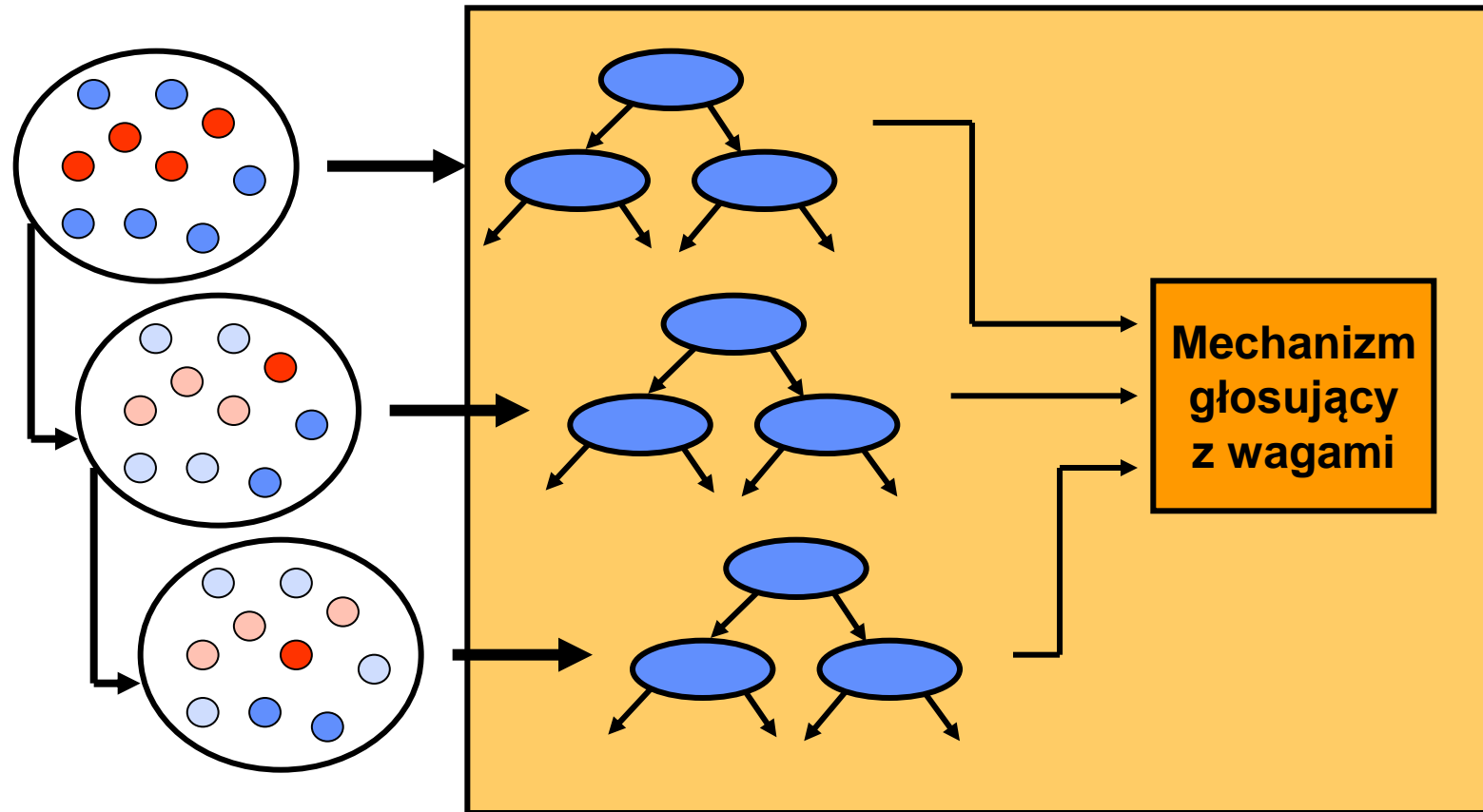
- W celu poprawy zdolności predykcyjnych można spróbować połączenia kilku klasyfikatorów w jeden,
- Istnieje kilka metod łączenia klasyfikatorów.

Bagging



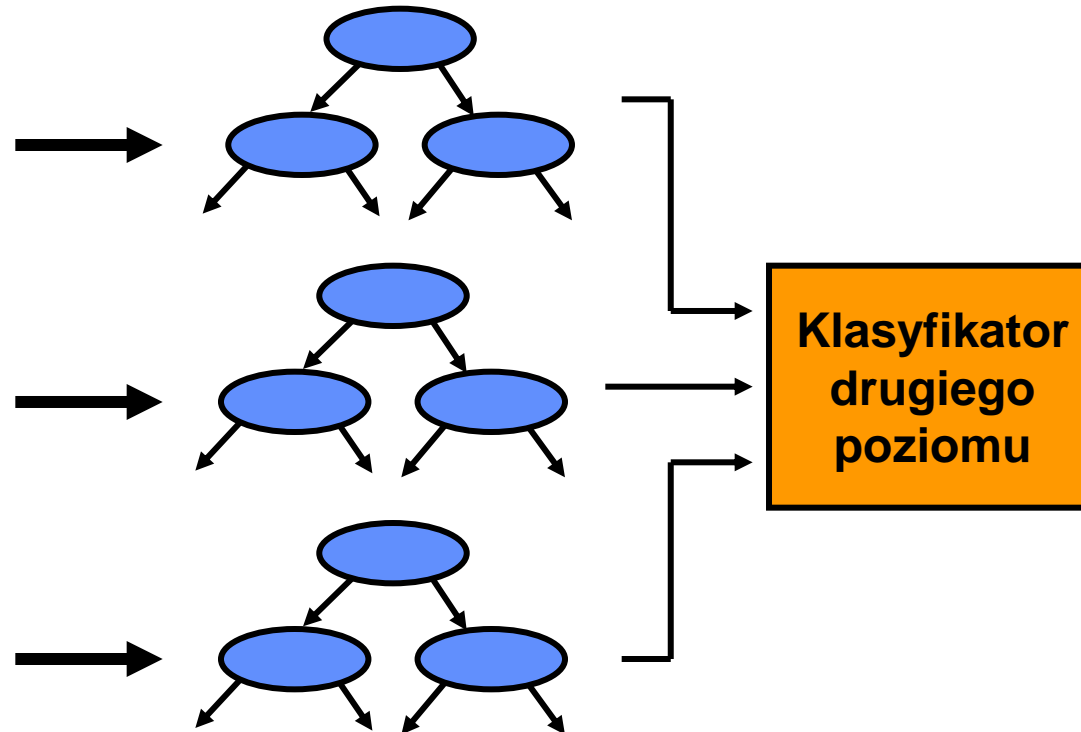
1. Z jednego zbioru przykładów tworzymy n zbiorów na zasadzie bootstrap
2. Na podstawie każdego zbioru tworzymy klasyfikator cząstkowy
3. Klasyfikator zagregowany tworzymy, dodając mechanizm głosujący

Boosting



1. Budujemy klasyfikator na podstawie zbioru uczącego
2. Przykładom źle sklasyfikowanym zwiększamy wagę; budujemy klasyfikator
3. Iteracyjnie powtarzamy krok 2 do osiągnięcia warunku stopu
4. Dodajemy mechanizm głosujący z wagami

Stacking



1. Tworzymy n klasyfikatorów

2. Dodajemy na ich wyjściu klasyfikator drugiego poziomu, który jest odpowiedzialny za decyzję, któremu klasyfikatorowi cząstkowemu zaufać

3. Istnieje wiele wariantów tej metody

Dziękujemy za uwagę

Zapraszamy na wykład:

KLASTERYZACJA I SEGMENTACJA cz. 1