

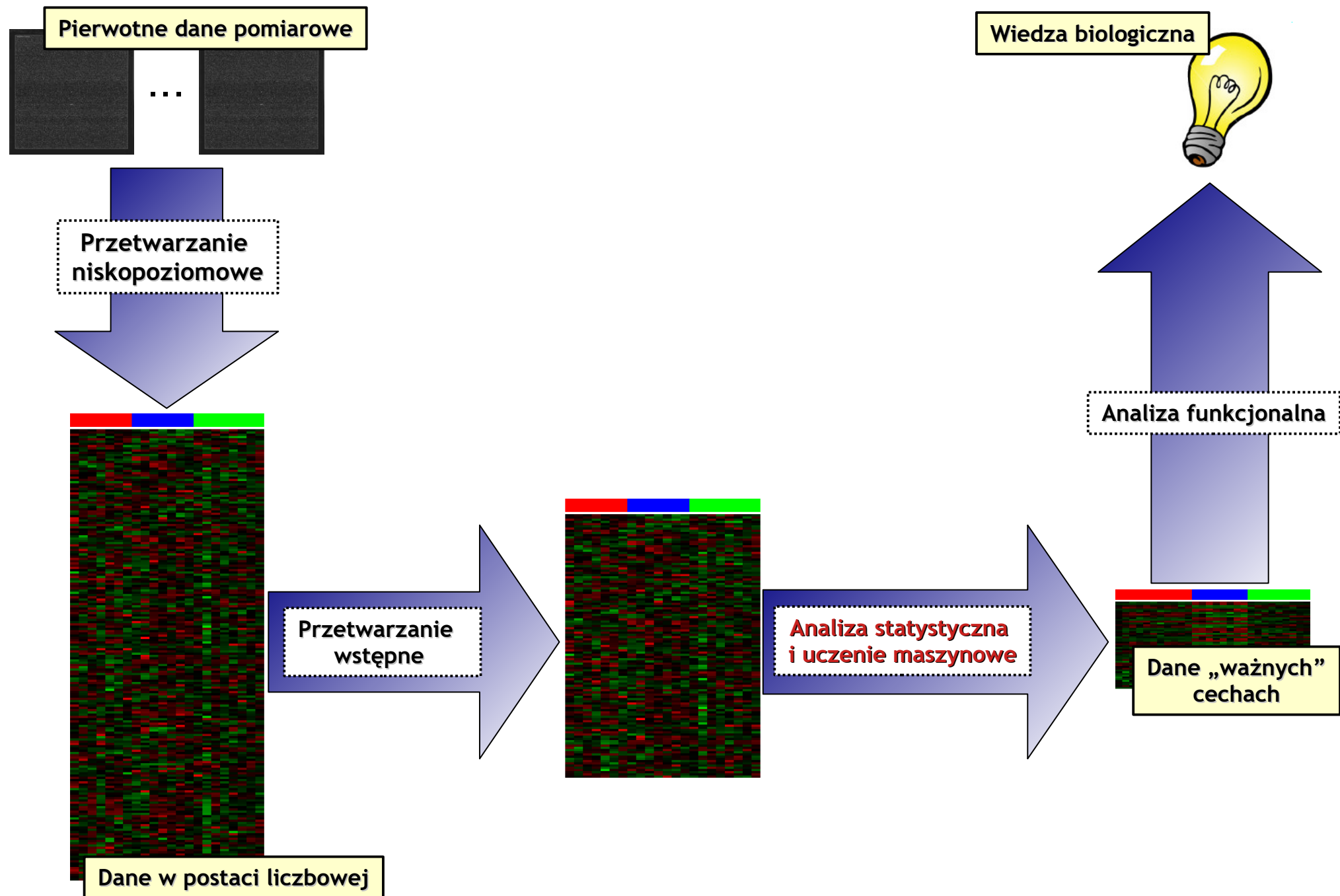
Uczenie maszynowe w bioinformatyce

Wykład 5: metody uczenia się maszyn

Tymon Rubel

**Zakład Elektroniki Jądrowej i Medycznej
Instytut Radioelektroniki i Technik Multimedialnych PW**

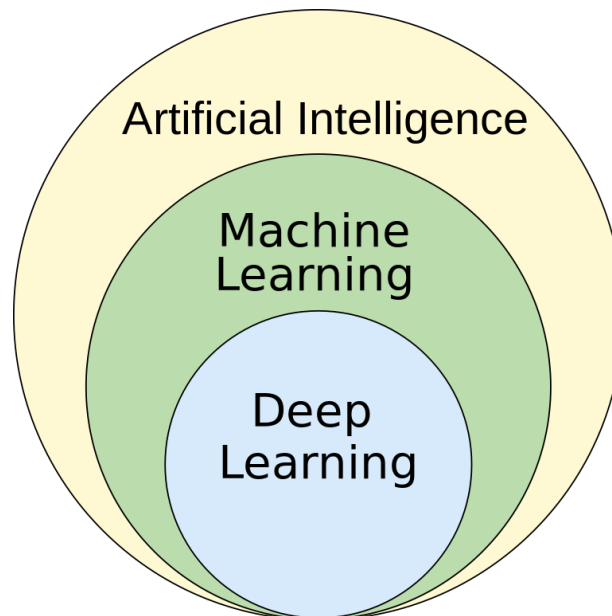
Etapy przetwarzania i analizy danych



Uczenie się maszyn

Uczenie się maszyn lub **uczenie maszynowe (ML – Machine Learning)** jest dziedziną nauki zajmującą się tworzeniem modeli matematycznych i algorytmów, które potrafią samodzielnie zmieniać swoją strukturę i/lub wartości parametrów w sposób zależny od **informacji zawartych w zbiorze składającym się z przykładów uczących (zbiorze uczącym)** lub na podstawie **interakcji ze środowiskiem, w którym działają**. Modele te powinny cechować się zdolnością uogólniania nabytej wiedzy, czyli skutecznego działania na wcześniej nieprezentowanych im danych.

Tak zdefiniowane uczenie maszynowe mieści się w obszarze zagadnień związanych z problematyką sztucznej inteligencji. Jednocześnie jest również silnie powiązane ze współczesną statystyką obliczeniową i metodami optymalizacji globalnej.



Uczenie się maszyn

Techniki uczenia się maszyn typowo dzieli się na trzy ogólne grupy.

- **Metody działające bez nadzoru (nienadzorowane, *unsupervised learning*)** pozwalają prowadzić wnioskowanie w oparciu o same dane, bez korzystania z dodatkowych informacji, w tym też tych dotyczących wartości oczekiwanych na wyjściu modelu po podaniu na jego wejście przykładów uczących.

W ich przypadku przykładami uczącymi są po prostu wektory ze zbioru danych.

Służą one przede wszystkim do rozwiązywania zadania **klasteryzacji (grupowania)** oraz do redukcji **wymiarowości danych**.

Uczenie się maszyn

Techniki uczenia się maszyn typowo dzieli się na trzy ogólne grupy.

- **Metody działające pod nadzorem (nadzorowane, *supervised learning*)** tworzą model na podstawie danych oraz dodatkowej wiedzy, wynikającej np. ze znajomości oczekiwanych wartości wyjściowych lub sposobu przypisania badanych obiektów do wcześniej ustalonych klas.

Przykładami uczącymi dla technik działających pod nadzorem najczęściej są pary złożone z wektora cech opisującego obiekt oraz oczekiwanej wartości wyjściowej.

Przykładami zagadnień, w których używa się metod nadzorowanych są **klasyfikacja**, **regresja (predykcja)** oraz powiązana z nimi **selekcja cech**.

Techniki uczenia się maszyn typowo dzieli się na trzy ogólne grupy.

- **Techniki uczenia ze wzmocnieniem (*reinforcement learning*)** istotnie różnią się od dwóch pozostałych podejść tym, że nie korzystają w przykładów uczących. Zamiast tego trening odbywa się poprzez interakcję ze środowiskiem, od którego model otrzymuje informacje o aktualnym stanie i ocenę swojego działania. Celem procesu nauki jest maksymalizacja oceny uzyskiwanej przez model.

Tego typu metody znajdują głównie zastosowanie przy rozwiązywaniu problemów związanych z użyciem sztucznej inteligencji do grania w różnego rodzaju gry (np. szachy, ale również na giełdzie) lub sterowania systemami autonomicznymi.