

Czym jest uczenie maszynowe?

- Automatyzacja automatyzacji
- Programowanie programów programujących programy
- Znajdowanie zależności w zbiorach danych
- Dopasowanie krzywych

Tradycyjne programowanie



Uczenie maszynowe



Ile jest algorytmów uczenia maszynowego?

- Dziesiątki tysięcy
- Setki nowych każdego roku
- Każdy składa się z trzech zasadniczych elementów
 - **Reprezentacja**
 - **Ewaluacja**
 - **Optymalizacja**

Reprezentacja

- Drzewa decyzyjne
- Układy oparte na regułach
- Modele graficzne (Bayes, sieci Markova)
- Sieci Neuronowe
- Maszyna wektorów nośnych

Ewaluacja

- Zgodność ze spodziewanymi wynikami
- Błąd kwadratowy
- Prawdopodobieństwo
- Funkcja kosztu
- Entropia
- Dywergencja K-L

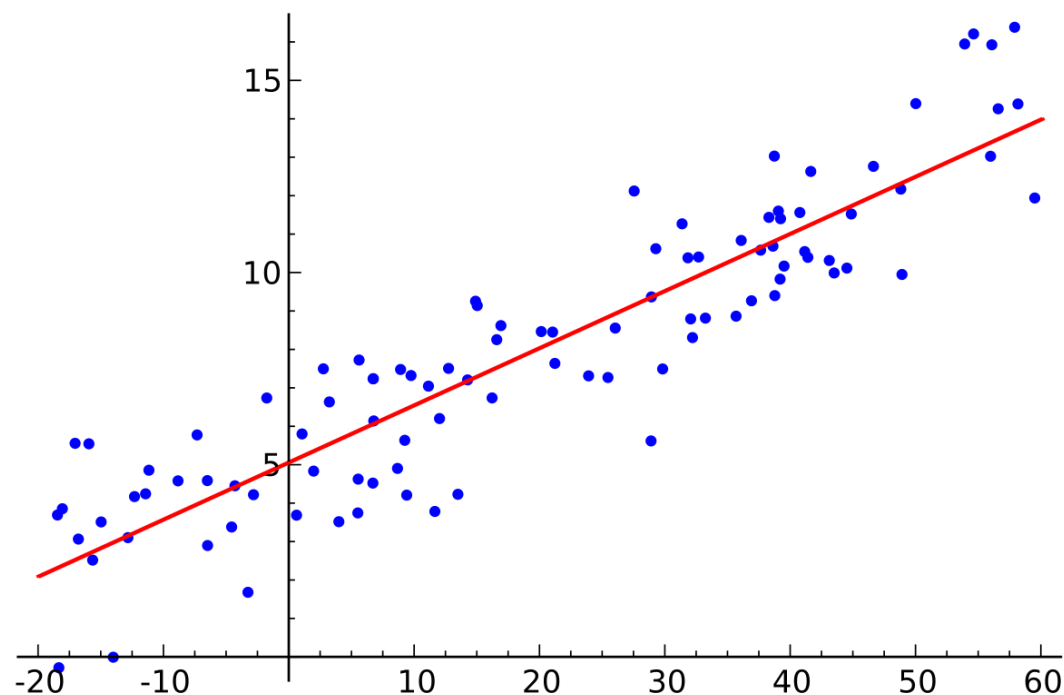
Optymalizacja

- Kombinatoryczna
 - Wyszukiwanie zachłanne
- Wypukła
 - Spadek gradientu
- Więzy
 - Programowanie Liniowe

Typy Uczenia Maszynowego

- Pod nadzorem
 - Wymaga zbioru wejść i odpowiadającym im wyjść
 - Np. Obrazki z podpisami
- Bez nadzoru
 - Tylko dane, brak wyjść
 - Algorytm znajduje zależności w danych
- Ze wsparciem
 - Tak jak z nadzorem tylko zamiast wyjść mamy funkcję nagrody/kary

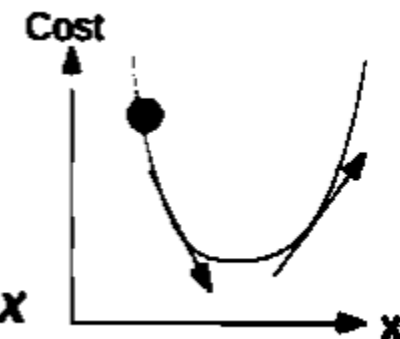
Regresja Liniowa



$$f = wx$$

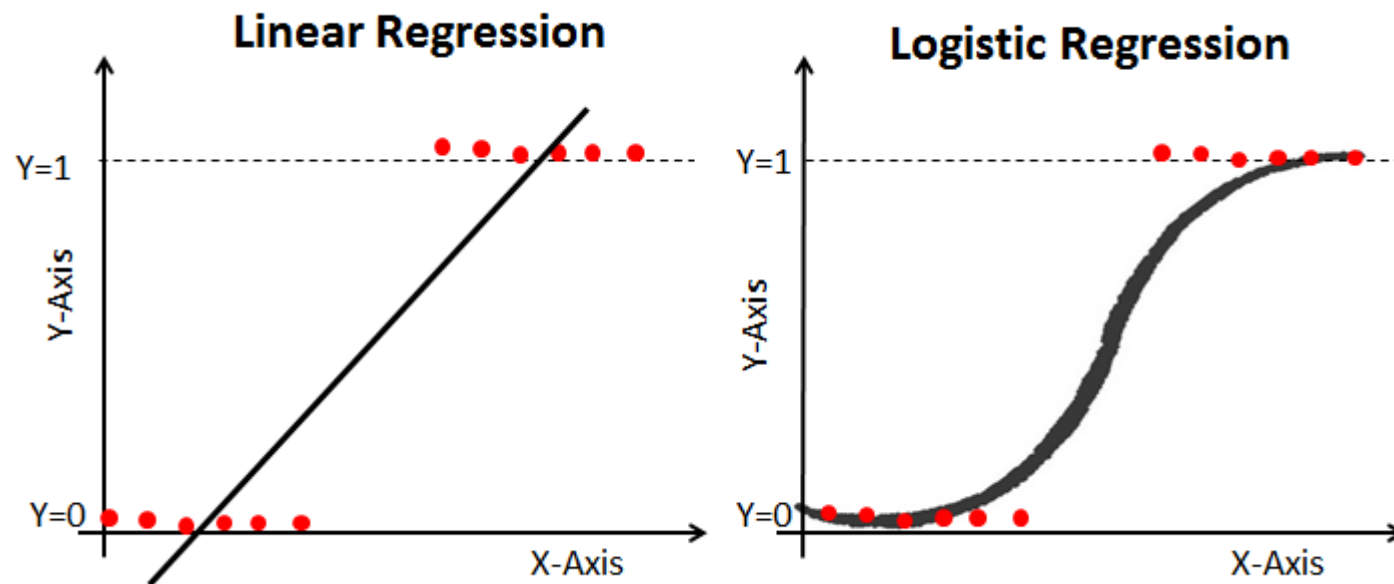
$$Cost \propto \sum (Y_{pred} - Y_{actual})^2$$

$$\nabla_w \propto \sum (Y_{pred} - Y_{actual})x$$



- Zmienne ciągłe (ceny akcji, ceny produktów)
- Funkcja kosztu jest proporcjonalna do sumy kwadratów różnic między danymi treningowymi i przewidywaniami
- Spadek gradientu do minimalizacji funkcji kosztu

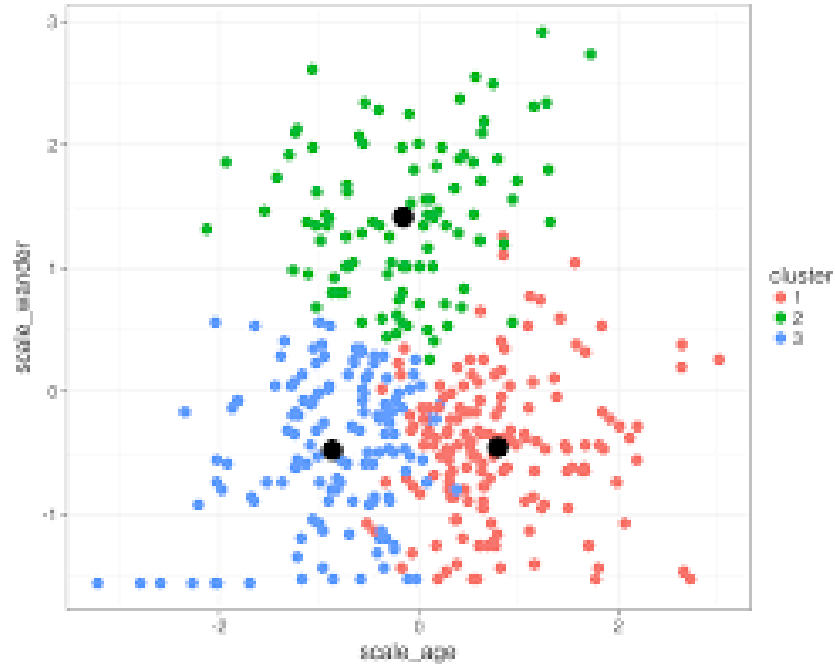
Regresja Logistyczna



- Przydatne przy klasyfikacji (spam/nie-spam, człowiek/nie-człowiek)
- Logarytmiczna funkcja kosztu
- Optymalizacja wag przy pomocy spadku gradientu

$$J(w) = \sum [-y \log(\Phi(z)) + (1-y) \log(1-\Phi(z))]$$

Algorytm Centroidów



$$d(x, r) = \sum_{j=1}^n (x_j - r_j)^2$$

- Kategorie podzielone na klastry.
- Algorytm sukcesywnie dopasowuje wektory kodowe do istniejących danych i w miarę potrzeb przesuwa błędnie zakwalifikowane wektory danych do innych grup.
- Problem stanowi początkowy wybór wektorów kodowych.

AI vs Szachy (chronologia)



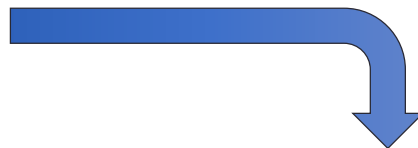
1912



1950



1970



1997



Dzisiaj



Dlaczego?

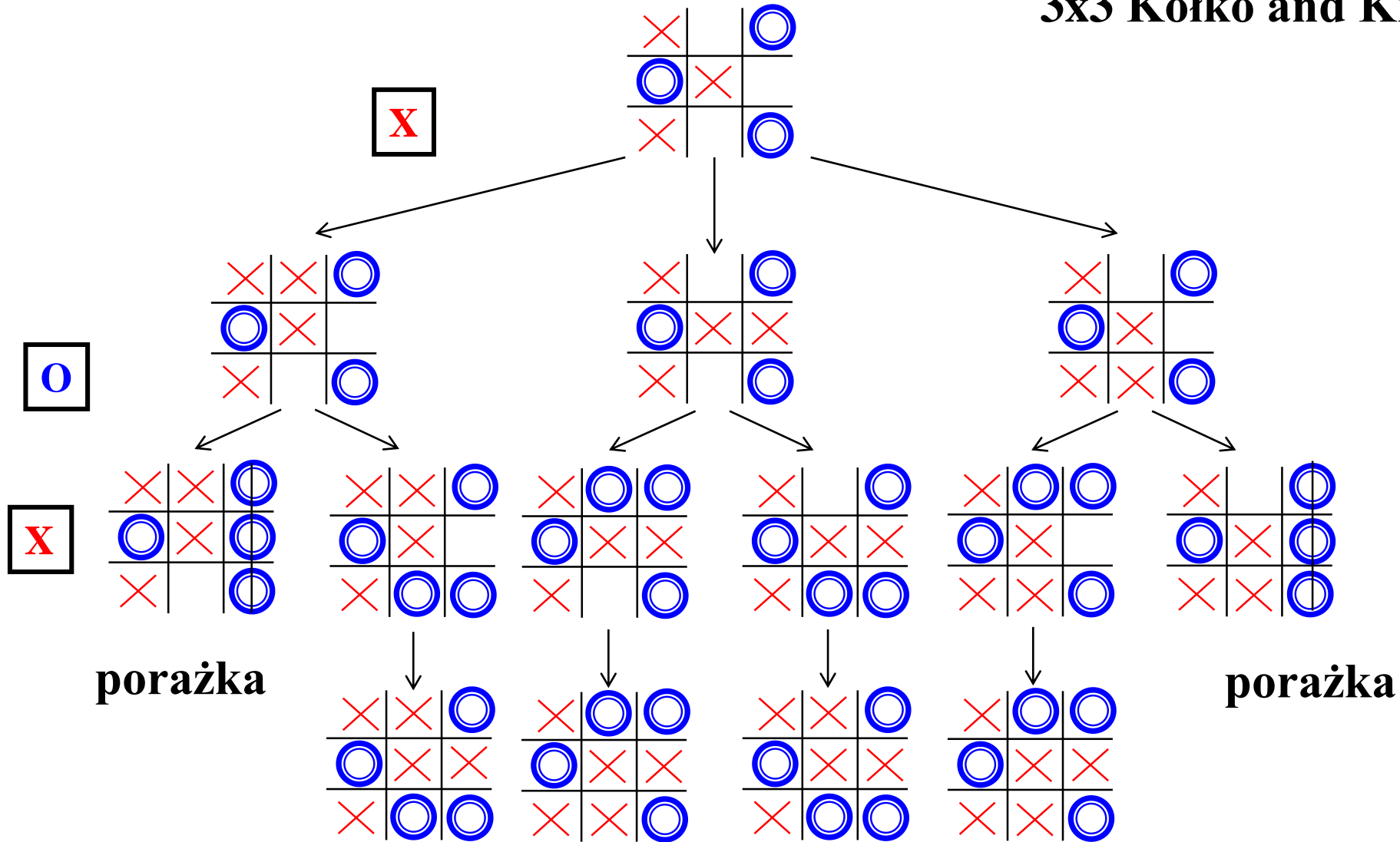
- Sprawność w grach strategicznych = inteligencja (jakaś część)
- Ocena sytuacji, planowanie i wyciąganie wniosków
- Podobieństwo do rozwiązywania rzeczywistych problemów
- Łatwo przenieść reguły na język zrozumiały dla komputera
- Co najmniej NP zupełne

Gry planszowe vs rzeczywistość

- Stała liczba graczy
- Skończona, najczęściej mała liczba graczy (np. 2)
- Niezmienne zasady
- Znane zasady
- Sekwencyjność
- Determinizm
- Pełna informacja o świecie

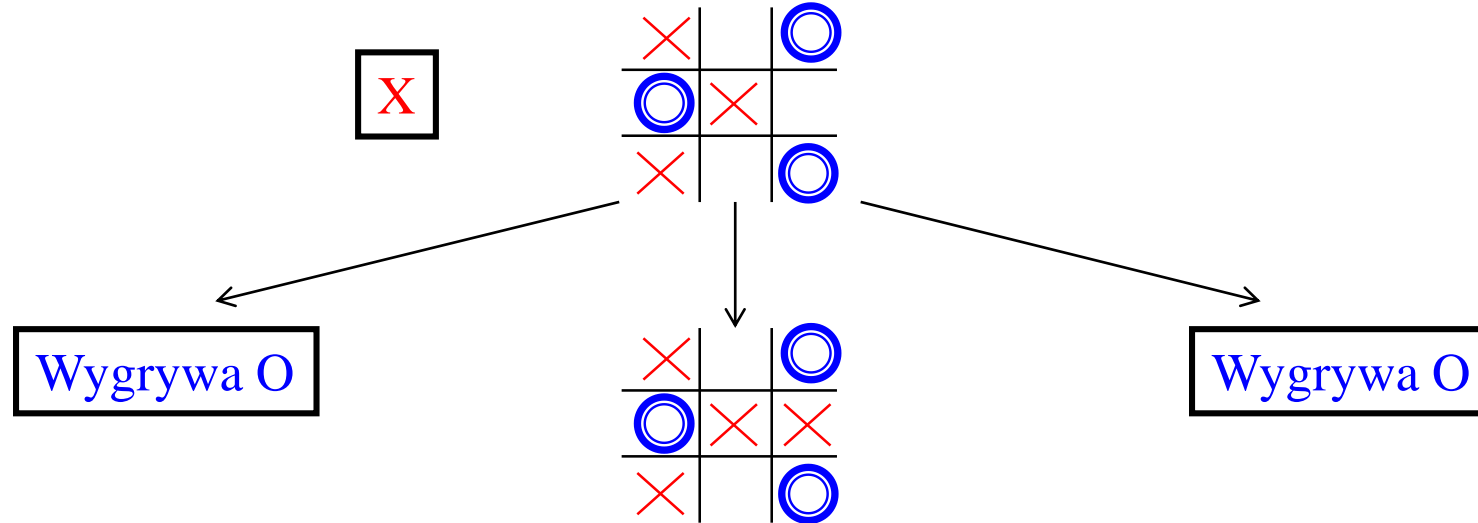


porażka



**Każda możliwa kombinacja ustawień
kółek i krzyżyków oceniana jest na -1, 0
lub 1 i ta wartość przechowywana jest w
wierzchołku drzewa decyzyjnego.**

Co jeśli przeciwnik nie gra optymalnie?



- Zaczynamy od dołu i idziemy w górę
- Używamy przechodzenia drzewa algorytmem DFS
- To jest MiniMax