

# Metody oparte na sztucznej inteligencji i analizie danych

By kama3 kama3  
lis 3, 2025

## klasyczne ML (ang. Machine Learning) i DL (ang. Deep Learning)

Znane są metody diagnostyki w kontekście istniejących metod takich jak klasyczne metody uczenia maszynowego ML (ang. Machine Learning) i DL (ang. Deep Learning). Są to rozwijane modele do klasyfikacji i prognozowania uszkodzeń na podstawie sygnałów (szczególnie drgań i przebiegów prądowych). Głębokie uczenie (Deep Learning) koncentruje się na wykorzystaniu wielowarstwowych sieci neuronowych do wykonywania zadań takich jak klasyfikacja, regresja i uczenie reprezentacji. Dziedzina ta czerpie inspirację z neurobiologii biologicznej i koncentruje się na układaniu sztucznych neuronów w warstwy i „trenowaniu” ich do przetwarzania danych. Zasadniczo, głębokie uczenie odnosi się do klasy algorytmów uczenia maszynowego (ang. Machine Learning), w którym hierarchia warstw służy do przekształcania danych wejściowych w coraz bardziej abstrakcyjną i złożoną reprezentację.

[Read more](#)

## SVM (ang. Support Vector Machine)

W uczeniu maszynowym maszyny wektorów nośnych (SVM), znane również jako sieci wektorów nośnych to nadzorowane modele o maksymalnym marginesie z powiązanymi algorytmami uczenia, które analizują dane pod kątem klasyfikacji i analizy regresji. Klasyfikator oparty na szukaniu hiperpowierzchni separujących klasy; bardzo skuteczny przy małej liczbie danych.

[Read more](#)

## k-NN (ang. k-Nearest Neighbors)

Klasyfikator k-NN klasyfikuje na podstawie "najbliższych sąsiadów" w przestrzeni cech. Jest to prosta metoda, ale kosztowna obliczeniowo w przypadku dużych zbiorów. Wadą algorytmu k-NN jest jego wrażliwość na lokalną strukturę danych. W klasyfikacji k-NN funkcja jest aproksymowana jedynie lokalnie, a wszelkie obliczenia są odkładane do momentu ewaluacji funkcji.

[Read more](#)

## Random Forest / Decision Trees

Losowe lasy decyzyjne to metoda uczenia zespołowego stosowana w klasyfikacji, regresji i innych zadaniach. Technika ta polega na konstruowaniu wielu drzew decyzyjnych podczas treningu, a końcowy wynik klasyfikacji jest określany przez klasę wybraną przez większość drzew, a w przypadku regresji przez średnią lub uśrednioną prognozę poszczególnych drzew. Modele drzew decyzyjnych lub ich zespoły są szybkie i dobrze działają na danych tablicowych.

[Read more](#)

## Naive Bayes

Klasyfikator probabilistyczny, zakładający niezależność cech; nadaje się do szybkich prototypów. W statystyce naiwne klasyfikatory Bayesa to rodzina „klasyfikatorów probabilistycznych”, która zakłada, że cechy są warunkowo niezależne, biorąc pod uwagę klasę docelową. Innymi słowy, naiwny model Bayesa zakłada, że informacja o klasie dostarczana przez każdą zmienną nie jest powiązana z informacją pochodzącą z pozostałych, a predyktory nie dzielą się żadną informacją.

[Read more](#)

## XGBoost / LightGBM

LightGBM, skrót od Light Gradient Boosting Machine, to otwartoźródłowy rozproszony framework gradientowego wzmacniania dla uczenia maszynowego. Bazuje na algorytmach drzew decyzyjnych i jest wykorzystywany do rankingowania, klasyfikacji i innych zadań uczenia maszynowego. Nacisk w rozwoju położono na wydajność i skalowalność.

XGBoost to skalowalna, wysokowydajna biblioteka do uczenia maszynowego oparta na algorytmie ekstremalnego wzmacniania gradientowego (Extreme Gradient Boosting), który buduje model iteracyjnie, dodając nowe drzewa decyzyjne, aby korygować błędy poprzednich.

XGBoost buduje swoje drzewa w sposób „poziomowy”, a LightGBM „liściowy”. Podejście liściowe dzieli liść na podstawie najlepszej globalnej opcji – czyli podziału opartego na najlepszej globalnej stracie w porównaniu z najlepszą stratą gałęzi.

LGBM jest znacznie bardziej wydajny obliczeniowo i bardziej podatny na przeuczenie.

**Najważniejsze zalety** XGBoost widoczne w implementacji:

Łatwa integracja z bibliotekami scikit-learn.

Możliwość pracy z różnymi typami danych (klasyfikacja binarna, wieloklasowa, regresja)

Wbudowana obsługa brakujących wartości.

Wysoka wydajność nawet przy domyślnych ustawieniach.

Zoptymalizowane wersje algorytmów gradient boosting; skuteczne i dokładne w klasyfikacji sygnałów.

[Read more](#)

## Logistic Regression

Prosty klasyfikator binarny; rzadko używany samodzielnie, ale czasem jako część większego pipeline'u. Regresja logistyczna jest jednocześnie bardzo skutecznym algorytmem klasyfikacji. W statystyce model logistyczny (lub model logitowy) to model statystyczny, który modeluje logarytm szans zdarzenia jako kombinację liniową jednej lub więcej zmiennych niezależnych. W analizie regresji regresja logistyczna (lub regresja logitowa) szacuje parametry modelu logistycznego (współczynniki w kombinacjach liniowych lub nieliniowych).

[Read more](#)