

# Zajęcia 14. Klasyfikacja

Wybierz z powyższej listy dwa dowolne klasyfikatory (inne niż SVM), wyszukaj informacje na ich temat i krótko opisz:

- 1) Jak działa,
- 2) Dlaczego warto byłoby go przetestować w zadaniu klasyfikacji recenzji/opinii klientów.

Decision Tree (Drzewo decyzyjne)

- 1) Mamy  $n$ -wymiarową przestrzeń i za pomocą algorytmu dzielimy ją na pewną liczbę "koszyków", które z kolei pozwalają nam powiedzieć, jak ocenić  $n+1$  decyzję (wpada do danego koszyka po tym, jak odpowiadamy na drzewku)
- 2) Jest bardzo użyteczne do klasyfikowania - wystarczy stworzyć kolejne pytania dotyczące danej opinii (np. czy zawiera słowo X) i na podstawie odpowiedzi stworzyć graf, który da nam ostateczną odpowiedź.

Logistic Regression (Regresja logistyczna)

- 1) Wykorzystywana jest do oceny prawdopodobieństwa w rozkładzie dwumianowym, oceniana na podstawie danych i dająca nam szansę na to, że następna dana będzie 0 lub 1 (pozytywna lub negatywna w naszym wypadku)
- 2) Jeżeli mamy już dużo opinii, to możemy dopasować dane, które dla kolejnej opinii dadzą nam szansę na to, że jest ona pozytywna lub negatywna. W przeciwieństwie do zwykłej regresji liniowej, daje nam to większą kontrolę nad false positives i false negatives.