Supervised Learning

*To podgrupa algorytmów Machine Learning, które uczą się przy wsparciu – potrzebują nie tylko danych wejściowych, ale też danych wynikowych, których dostarczanie w konsekwencji ‘uczy’ algorytm rozpoznawać wzór i przez to zwracać oczekiwany przez nas efekt. Najlepiej wykorzystywać jeśli mamy do czynienia z danymi, które posiadają jakieś etykiety.*

*Zadania, w których sprawdza się SL, to:*

* Klasyfikacja – czyli zbieranie dostarczanych danych w pewne zbiory na podstawie jasnych wytycznych, np.:
  + Decydowanie, na podstawie przedstawionych objawów, czy dany pacjent spełnia kryteria potrzebne do postawienia konkretnej diagnozy
  + Badanie zachowań konsumenckich – przypisywanie do kategorii np. pod kątem częstotliwości zakupów
  + Rozpoznawanie ptaków – są aplikacje poświęcone rozpoznawaniu in real time głosów ptaków na podstawie wcześniej wgranych nagrań przypisanych do konkretnych gatunków
* Regresja – przewidywanie kolejnych elementów ciągu, co ważne, elementów liczbowych, na podstawie cech, np
  + Badanie zmian klimatycznych, tworzenie modeli temperaturowych, np. zmian temperatur oceanów
  + Analizowanie skuteczności kampanii reklamowych
  + Przewidywanie skuteczności działania terapii

Unsupervised Learning

*To rodzaj Machine Learningu, który nie potrzebuje – jak supervised – etykiet, opisów danych. Wymaga jedynie dostarczenia danych wejściowych, które następnie sam przetwarza identyfikując wzorce, zależności itp.*

* + Clustering – przypisywanie na podstawie wielu cech do większych grup
    - Pierwsze co mi przychodzi do głowy to świeży przykład targetowania wyborców podczas kampanii na podstawie demografii, aktywnościach czy lokalizacji, np. skupianie się na konkretnych okręgach wyborczych, które dają szanse na zmianę decyzji dotyczących głosowania
    - Bardzo ciekawy dla mnie przykład, który znalazłam dotyczy statystyk sportowych – na podstawie danych o zachowaniach graczy można ich pogrupować i dostosowywać treningi do konkretnych grup
    - Genetyka – np. grupowanie sekwencji genów, wg. Ich ekspresji
  + Dimension Reduction – redukowanie danych, które mogą nie wnosić wartości do analizy, a wręcz ją utrudniać
    - Analizowanie ryzyka udzielania kredytów – zawężanie danych o klientach do tych najważniejszych czynników, które mają największy wpływ na poziom ryzyka
    - Algorytmy rozpoznające twarze na zdjęciach – dzięki DR pomijają dużą liczbę nieistotnych informacji, pikseli, które nie są kluczowe dla rozpoznania danej osoby
    - Ciekawy przykład – kompresowanie nagrań z monitoringu zmierzająca do wybrania kluczowych klatek