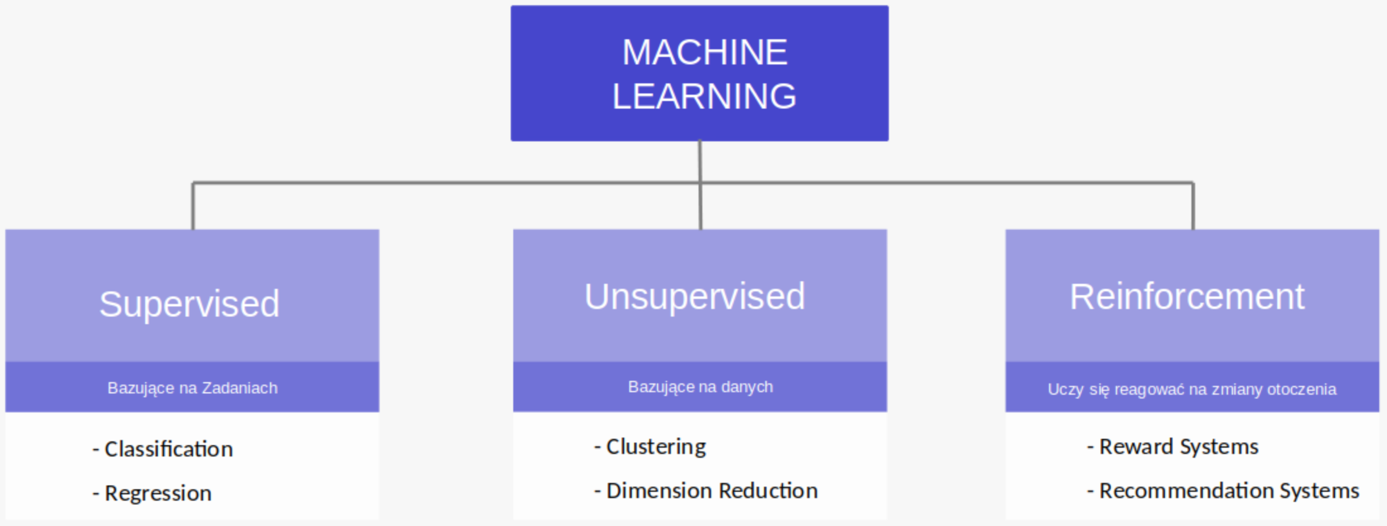
TASK 12.2

# MACHINE LEARNING

Machine Learning (ML) czyli uczenie maszynowe to część sztucznej inteligencji (AI) i informatyki, która skupia się na wykorzystaniu danych i algorytmów do naśladowania sposobu, w jaki uczą się ludzie – stopniowo zwiększając swoją dokładność. Dzięki zastosowaniu metod statystycznych algorytmy są „szkolone” w celu klasyfikowania, przewidywania i odkrywania kluczowych zależności w projektach opartych na przetwarzaniu danych. Ich interpretacja jest bowiem wykorzystywana do znajdowania odpowiedzi na najistotniejsze pytania biznesowe.

Wyróżniamy kilka rodzajów algorytmów ML. Na schemacie poniżej przedstawiamy ich typy.



## Supervised Learning

Supervised Learning (SL), nazywane uczeniem nadzorowanym, znane również jako nadzorowane uczenie maszynowe, jest podkategorią uczenia maszynowego i sztucznej inteligencji. Jest definiowany przez użycie oznaczonych zestawów danych do trenowania algorytmów, które dokładnie klasyfikują dane lub przewidują wyniki. Gdy dane wejściowe są wprowadzane do modelu, dostosowuje on swoje wagi, aż model zostanie odpowiednio dopasowany. Uczenie nadzorowane pomaga organizacjom rozwiązywać różne rzeczywiste problemy na dużą skalę, np. klasyfikowanie spamu w osobnym folderze ze skrzynki odbiorczej. W tym rodzaju ML wyróżniamy dwa typy: klasyfikacje i regresję.

### Klasyfikacja

Klasyfikacja to proces znajdowania funkcji, która pomaga w podziale zbioru danych na klasy w oparciu o różne parametry. W klasyfikacji program komputerowy jest szkolony na zbiorze danych szkoleniowych i na podstawie tego szkolenia kategoryzuje dane na różne klasy. Przykłady:

* Podany powyżej przykład klasyfikacji maili (spam / nie spam)
* Określanie rodzaju znaku drogowego na podstawie zdjęcia (znak informacyjny / nakazu / zakazu itp.)
* Określanie płci osoby na zdjęciu (kobieta / mężczyzna)

### Regresja

Regresja to proces znajdowania korelacji między zmiennymi zależnymi i niezależnymi. Pomaga w przewidywaniu zmiennych ciągłych, takich jak przewidywanie trendów rynkowych, przewidywanie cen domów itp. Przykłady:

* Przewidywania cen nieruchomości
* Przewidywanie zużycia energii
* Przewidywanie pogody

## Unsupervised Learning

Unsupervised Learning (UL) to uczenie bez nadzoru, znane również jako uczenie maszynowe bez nadzoru, wykorzystuje algorytmy uczenia maszynowego do analizowania i grupowania nieoznakowanych zestawów danych. Algorytmy te odkrywają ukryte wzorce lub grupowanie danych bez potrzeby interwencji człowieka. Jego zdolność do odkrywania podobieństw i różnic w informacjach czyni go idealnym rozwiązaniem do eksploracyjnej analizy danych, strategii cross-sellingu, segmentacji klientów i rozpoznawania obrazu.

### Clustering

Clustering to nienadzorowana metoda uczenia maszynowego służąca do identyfikowania i grupowania podobnych punktów danych w większych zbiorach danych bez troski o konkretny wynik. Grupowanie (czasami nazywane analizą skupień) jest zwykle używane do klasyfikowania danych w struktury, które są łatwiejsze do zrozumienia i manipulowania. Przykłady:

* Analiza danych z social mediów
* Analiza rynku pod kątem wyboru klientów
* Grupowanie danych biomedycznych

### Dimension Reduction

Redukcja wymiarowości to uczenie maszynowe (ML) lub technika statystyczna zmniejszania liczby zmiennych losowych w problemie poprzez uzyskanie zestawu głównych zmiennych. Przykłady:

* Analiza danych tekstowych na podstawie jakiś fragmentów (np. fraz)
* Analiza wykrywania „włamań” w sieci na podstawie aktywności użytkowników (które aktywności mają wpływ, a które nie)
* Analiza ekspresji genów