**Uczenie maszynowe**

Zasady ćwiczeń:

1. Format rozwiązania: notebook (.ipynb). Widzimy kod, rozwiązanie, a jak trzeba grafikę.
2. Wykonane skrypty wkładamy do Zadania na platformie Moodle (czas jest ograniczony).
3. Na zajęciach należy zreferować rozwiązania (**proszę, żeby nie było sytuacji, że nikt nie referuje**).

Lab: **7** Uczenie nienadzorowane + inżynieria cech + metody zespołowe

Jeśli metodykę któregoś z etapów pracy można uogólnić, taki etap proszę zakodować w postaci metody  
(z ew. parametrami).

Proszę opracować paradygmaty uczenia nienadzorowanego:

1. klasteryzacja
2. szacowanie gęstości rozkładu (kernel density estimation)
3. detekcja wartości odstających

Inżynieria cech:

1. metody selekcji cech
2. zmniejszanie wymiarowości danych

Proszę opracować paradygmaty uczenia zespołowego:

1. bagging
2. boosting
3. lasy losowe

W ramach danego paradygmatu uczenia przedstawić:

* podstawy teoretyczne (krótko)
* metody/typy
* parametry, regularyzację, optymalizację
* implementację (własną lub biblioteczną): uczenie, predykcja (testowanie)
* typowe metryki wraz z metodami ich wyliczania
* wizualizację: zbioru/modelu
* uczenie na wybranym zbiorze danych (przy wielu metodach w ramach paradygmatu, na różnych zbiorach danych)

Najważniejszą częścią wykonanej pracy jest prezentacja i dyskusja na ćwiczeniach.

Przykładowe archiwa danych:

www.kaggle.com/datasets

archive.ics.uci.edu/ml

datasetsearch.research.google.com