

Projekt 1

1. Zbiór danych: TNG, ok 18000 próbek, 20 klas. Wykorzystujemy gotowe dane reprezentujące tekst blogów w postaci wektorów (dostarcza prowadzący)
2. Celem projektu jest porównanie dwóch klasyfikatorów: klasyfikatora AdaBoost użytego dla pełnych danych oraz sieci neuronowej MLP.
3. Ocenic moc obu (*robustness*) klasyfikatorów na ich jakość w zależności od stopnia zaszumienia wejściowych danych.
4. Jak zmieni się jakość klasyfikatorów po użyciu transformacji PCA.
5. Wyniki jakości klasyfikatorów oceniać na bazie krzyżowej-walidacji, (accuracy - Loss-błąd, Krzywa ROC, Precision-Recall, (pole pod krzywymi) F1).
6. Dokonać porównania wyników tzn: accuracy i loss. Jakość klasyfikatorów porównywać dla 3 budżetów czasowych T1, T2, T3 (krótki, średni i długi).
7. Należy dokonać badań:
 - a. 5,6. **(3.5)**
 - b. 3,4 **(4.0)**
 - c. Znaleźć zestaw cech, który jest głównie odpowiedzialny za przynależność do odpowiednich klas. **(5.0)**
8. Należy sporządzić raport z projektu.

W nawiasie ocena maksymalna.